Vietnam National University - Ho Chi Minh City
University of Science
Faculty of Mathematics and Computer Science

Linear control systems and the regularity of the minimum time function

BACHELOR'S THESIS

STUDENT: Tu Nguyen Thai Son
ID: 1111279

ADVISOR: Khai .T Nguyen
PSU Mathematics Department

May 31, 2015

Linear Control Systems and the regularity of the minimum time function

Student: Tu Nguyen Thai Son

May 31, 2015

Abstract

Consider the linear control system $x'(t) = Ax(t) + Bu(t)$ with the target $\{0\}$. The function $T(x)$ is the minimum time needed to steer a point x to $\{0\}$. It is well-known that under the Rank-Kalman condition, T is just continuous in the reachable set. Our main goal is to study the regularity of $T(x)$. In this thesis, we prove that T is a.e. twice differentiable. Moreover, we also obtain an explicitly formula of the set of singularities where the optimal control changes its sign.

Acknowledgements

First of all, I would like to express my sincere gratitude to my advisor Dr. Khai Tien Nguyen, for the continuous support my study and research. Thanks to him for his patience, motivation, enthusiasm, and immense knowledge. His guidance helped me in all the time of researching and writing my thesis.

I also wish to express my sincere thanks to Professor Trong Duc Dang, Dean of the Faculty, for the continuous encouragement. He helped me have a change to meet my advisor.

I am also very grateful to Dr. Vu Quang Huynh, my professor in the Faculty of Mathematics and Computer Science University of Science. I am extremely thankful and indebted to him for sharing expertise, sincere and valuable guidance as well as encouragement extended to me, for the last 4 years and during the preparation of my thesis.

Lastly, I would like to thank my family: my parents for giving birth to me at the first place, as well as my sister, for supporting me throughout my life.

Contents

1 Introduction

Consider the linear control system

$$\begin{cases} x'(t) &= Ax(t) + Bu(t) \qquad t > 0 \\ x(0) &= x_0 \end{cases} \tag{1.1}$$

where $A \in \mathbb{M}^{n\times n}$, $B \in \mathbb{M}^{n\times m}$, $1 \le m \le n$ and $U = [-1,1]^m \subset \mathbb{R}^m$.

- $U = [-1,1]^m \subset \mathbb{R}^m$ is the control set.
- $u : [0,\infty) \longrightarrow U$ is a control function.

We will write

$$x(t) = \begin{pmatrix} x^1(t) \\ x^2(t) \\ \vdots \\ x^n(t) \end{pmatrix} \qquad \text{and} \qquad u(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_m(t) \end{pmatrix}$$

The set of admissible control is

$$\mathcal{U}_{ad} = \{u : [0,\infty) \longrightarrow U : u \text{ is measurable}\}$$

A solution of (1.1) depends on initial state x_0 and the choice of admissible control u, denote a solution by $y^{x_0,u}(.)$ be the trajectory starting from x_0 associated with the control u.

Given a target set $S = \{0\} \subset \mathbb{R}^n$, we want to consider some basis questions

- Does exists a control $u \in \mathcal{U}_{ad}$ such that the system (1.1) can be steer to $S = \{0\}$ in a finite amount of time? Which initial point x_0 admit such a control like that?
- If there exist such a control like this, we call an control u is optimal, which is the control that steer x_0 to the target set S in the minimum time. "Does an optimal control exist?".
- How can we characterize an optimal control?
- How can we construct an optimal control?

In 2.3.3 and 2.3.4 we will give some examples about this system. Furthermore, under some conditions, we can prove that there exists the minimum time function

$$\begin{aligned} T : \mathbb{R}^n &\longrightarrow \mathbb{R} \\ t &\longmapsto T(x) = \inf_{u\in\mathcal{U}_{ad}} \{t > 0 : y^{x_0,u}(t) = 0\} \end{aligned}$$

Later, we know that T is continuous, but in general it's not smooth or even Lipschitz. Several papers were devoted to the partial regularity of T. In particular, we quote results devoted to establishing (semi) convexity/concavity properties of T under various assumptions (see [11], [12], [3], [6].)

The last part of this document concerns the regularity of T. In case

$$\text{dom}(T) = \{x \in \mathbb{R}^n : T(x) < \infty\} = \mathbb{R}^n$$

the main goal is finding the structure of the set of non-Lipschitz points

$$\mathcal{S} = \{x \in \mathbb{R}^n : T \text{ is not Lipschitz at } x\}$$

by using some tools in non-smooth analysis. By introducing the minimized Hamiltonian

$$H(x,\zeta) = \langle Ax, \zeta\rangle + \min_{a\in U}\langle Bu, \zeta\rangle$$

In 5 we will see that when $\zeta \in N_{\mathcal{R}(T(x))}(x)$, then

$$H(x,\zeta) \leq 0$$

Since the minimized Hamiltonian is constant and non-positive along every optimal trajectory, it is natural to expect that non-Lipschitz points of T lie exactly where such Hamiltonian vanishes. In 5 we will prove this characterization. We will use the result in [3] and [6], saying that the epigraph of T has locally positive reach.

Let $\mathcal{S}$ be the set of non-Lipschitz points of T, in 5 we will give the explicit formula of $\mathcal{S}$, using points which belong to an optimal pair (i.e., an optimal trajectory together with a corresponding adjoint arc) of (1.1) with vanishing Hamiltonian.

Finally, the positive reach property of epi(T) implies that T is locally semi-convex outside the closed set $\mathcal{S}$ (see theorem 5.1 in [7]). The structure of singularities of semi-convex functions is well understood (see chapter 4 in [12]). And $\mathcal{S}$ has Lebesgue measure 0, this implies T is a.e twice differentiable in $\mathbb{R}^n$.

1.1 Notations

We consider the Euclid space $(\mathbb{R}^n, \|.\|)$ denote all vector $x = (x_1, x_2, \dots, x_n)$ where $x_i \in \mathbb{R}$ for all $i = \overline{1,n}$. The Euclid norm is defined as $\|x\| = \sqrt{|x_1|^2 + \ldots + |x_n|^2}$.
As usual, the open ball with center $x \in \mathbb{R}^n$ and radius r, where $r > 0$ is defined as $B(x,r) = \{y \in \mathbb{R}^n : \|y - x\| < r\}$, the closed ball with center $x \in \mathbb{R}^n$ and radius r, where $r > 0$ is the closure of $B(x,r)$, and the sphere with center $x \in \mathbb{R}^n$ and radius r, where $r > 0$ is the boundary of $B(x,r)$, i.e $S(x,r) = \{y \in \mathbb{R}^n : \|y-x\| = r\}$. We also denote the scalar product is denoted by $\langle x, y\rangle$ or $x \cdot y$. Sometime, the Lebesgue measure on $\mathbb{R}^n$ is denoted by μ or $\mathcal{L}^n$.

In this document, we will use following notations when dealing to the linear optimal control (1.1) for convenient

- The set of admissible controls is

$$\mathcal{U}_{ad} = \{u : [0,\infty) \longrightarrow U : u \text{ is measurable}\}$$

- A solution of (1.1) depends on initial state x_0 and the choice of admissible control u, denote a solution by $y^{x_0,u}(.)$ be the trajectory starting from x_0 associated with the control u. (We will see (1.1) has unique solution which respect to u and x_0 later.)
- $U \subset \mathbb{R}^m$ is the control set, where $U = [-1,1]^m$
- $u : [0,\infty) \longrightarrow U$ is a measurable control function.

1.2 Structure of thesis

This thesis conclude of 5 sections.

1. Introduction. We introduce the problem, notations which will be used in through this thesis.

2. Preliminaries. We recall some background knowledge in function analysis, measure theory, first order original differential equation (ODE), viscosity,..., especially convex and non-convex, non-smooth analysis, which is the most important tool used to study the regularity of the minimum time function. In this section, we also proving almost basic properties of the solution (trajectory) of linear system. The dynamic programming principle is crucial for investigating lots of properties of the minimum time function later.

3. Controllability. We study when the linear system satisfies the small time controllable or full time controllable property. From this we will describe the reachable set $\mathcal{R}$ of all point which can be steered into the origin in finite time. In the end of this section, we can see that the minimum time function T is continuous from the reachable set $\mathcal{R}$ onto $\mathbb{R}$. Furthermore, under the Rank-Kalman's condition and an assumption of the matrix A, $\mathcal{R} = \mathbb{R}^n$ and thus T can be defined on whole space $\mathbb{R}^n$.

4. Optimal controls. This section concern when there exists an optimal control steers a given point (belong to reachable set) into the origin in the minimum amount of time. We will prove that such an optimal control is always existed, and furthermore it's unique by bang-bang principle. Furthermore we can describe the structure of thus control by maximum principle. Two examples will be given in this section.

5. The regularity of minimum time function. This is the main part of this thesis, we will claim that under some assumptions, the minimum time function is semi-convex in a dense subset of $\mathbb{R}^n$. We also describe the detailed structure of the set of non-Lipschitz point of this minimum time function.

2 Preliminaries

In this section, we assert some basic well–known knowledge of functional analysis, measure theory,...

2.1 Functional Analysis

2.1.1 Normed spaces

Let E be a vector space, a norm $\|.\|$ on E is a mapping $E \longrightarrow [0,+\infty)$ such that for every $x, y \in E$ and $\lambda \in \mathbb{R}$, $\|x\| = 0$ iff $x = 0$, $\|\lambda x\| = |\lambda|.\|x\|$ and $\|x + y\| \le \|x\| + \|y\|$.
Given a subset A in E, we call $\text{int}(A)$ denote the interior of A, $\overline{A}$ denote the closure of A and $\partial A = \overline{A} \cap \overline{X \backslash A}$ is the boundary of A.

A sequence $(x_n) \subset E$ is called **Cauchy** iff $\|x_n - x_m\| \longrightarrow 0$ as $m, n \longrightarrow \infty$. If every Cauchy sequence in E is convergence in E, we call E is a Banach space.

Let F be a normed space, a linear function $f : E \longrightarrow F$ is continuous iff it's continuous at 0, i.e there exists a constant $C > 0$ such that

$$\|f(x)\| \le C\|x\| \qquad \forall\, x \in X$$

And denote $\|f\| = \sup\{|f(x)| : \|x\| = 1\}$.

The set of all linear continuous function from E to F denoted by $L(E, F)$. Incase $F = \mathbb{R}$, we set $L(E, \mathbb{R}) = E^*$, which is called the **dual space** of E.

Consider a complex vector space E, an **inner product** is a map $\langle ., . \rangle : E \times E \longrightarrow \mathbb{C}$ s.t

(i) $\langle ax + by, z = a\langle x, z\rangle + b\langle y, z\rangle$ for all $x, y, z \in E$ and $a, b \in \mathbb{C}$

(ii) $\langle y, x\rangle = \overline{\langle x, y\rangle}$

(iii) $\langle x, x\rangle \in (0, \infty)$ for all nonzero $x \in E$.

We define $\|x\| = \sqrt{\langle x, x\rangle}$ for all $x \in E$, and also called this is a norm. And from the Schwarz Inequality $|\langle x, y\rangle| \le \|x\|.\|y\|$ yells that $(E, \|.\|)$ come to a normed space. If this space is complete, we call E is a Hilbert space which respect to the inner product $\langle ., . \rangle$.

An simple but very useful formula in Hilbert space is the Parallelogram Law, that is

Theorem 2.1 (Parallelogram Law): For all $x, y \in H$,

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2$$

In case E is a Hilbert space, we have a representation for linear continuous mapping

Theorem 2.2 (Riesze's theorem): If $f \in E^*$, there is a unique $y \in E$ such that $f(x) = \langle x, y\rangle$ for all $x \in E$.

So if E be a sub–vector space of $\mathbb{R}^n$, which is a normal Hilbert space with scalar product restricted on E, for any linear continuous function from E to $\mathbb{R}$, there exists a unique vector $\xi \in E$ such that

$$f(x) = \xi \cdot x = \langle \xi, x \rangle \qquad \forall\, x \in E$$

2.1.2 Weak topology

If E be normed space with dual space E^*, the strong topology on E is the topology generated by norm, weak topology on E, denote by $\sigma(E, E^*)$ is the topology generated by all linear continuous mapping from E to $\mathbb{R}$, i.e generated by E^*.

Similarly, the normed space E^* with the usual norm $\|f\| = \sup\{|f(x)| : \|x\| = 1\}$ has strong topology is the topology generated by this norm. The weak topology $\sigma(E^*, E^{**})$ is the topology generated by E^{**}. On the other hand, note that the canonical projection

$$\begin{aligned} \Phi : E &\longrightarrow E^{**} \\ x &\longmapsto \widehat{x} \qquad \text{where} \qquad \widehat{x}(f) = f(x) \end{aligned}$$

is a linear isometry, so this let's to identity $\widehat{x}$ with x and thus regard E^{**} as a superspace of E. In this case, we call the weak* topology $\sigma(E^*, E)$ on E^* is the topology generated by E. We have a basic fact about this topology, it's yell that $f_n \overset{*}{\rightharpoonup} f$ in the weak* topology $\sigma(E^*, E)$ iff $f_n(x) \longrightarrow f(x)$ for all $x \in E$, and

$$\|f\| \leq \liminf \|f_n\|$$

We state an important result

Theorem 2.3: [Banach–Alaoglu–Bourbaski theorem] If E is a normed vector space. The closed unit ball $B^* = \{f \in E^* : \|f\| \leq 1\}$ in E^* is compact in the weak* topology.

If f, g are two measurable functions from $[0, t] \longrightarrow \mathbb{R}^m$, and $f = g$ a.e, i.e the set $\{x : f(x) \neq g(x)\}$ has Lebesgue measure 0, then we identity $f \equiv g$ into a equivalent class $[f]$. We call $L^\infty([0, t], \mathbb{R}^m)$ be the set of all class measurable function from $[0, t] \longrightarrow \mathbb{R}^m$ s.t

$$\sup_{s \in [0,t]} |f(s)| < \infty$$

And define

$$\|f\|_\infty = \sup_{s \in [0,t]} |f(s)|$$

We recall that $L^\infty, \|.\|_\infty$ is a Banach space.
Similarly, $L^1([0, t], \mathbb{R}^m)$ be the set of all class measurable function from $[0, t] \longrightarrow \mathbb{R}^m$ s.t

$$\int_0^t |f(s)|\, ds < \infty$$

and define

$$\|f\|_{L^1} = \int_0^t |f(s)|\, ds$$

We recall that $L^\infty, \|.\|_\infty$ is a Banach space. And an important result

Theorem 2.4: We have $(L^1)^* = L^\infty$

So in this case, the weak* convergence in $\sigma(L^\infty, L^1)$ can be known as

$$f_n \xrightarrow{*} f \iff \int_0^t g(s)\cdot f_n(s)\,ds \longrightarrow \int_0^t g(s)\cdot f(s)\,ds$$

Finally, we have a useful weak* compactness theorem for L^∞, which is an corollary of Alaoglu theorem.

Theorem 2.5: Let's $(f_n) \subseteq L^\infty([0,t],\mathbb{R}^m)$ is a bounded sequence, then there exists a subsequence f_{n_k} and some $f \in L^\infty([0,t],\mathbb{R}^m)$ such that $f_{n_k} \xrightarrow{*} f$ in the weak* topology $\sigma(L^\infty, L^1)$.

Note that if we write $f = \left(f^{(1)}, f^{(2)}), \ldots, f^{(m)}\right)$, then clearly if $\|f_{n_k}^{(i)}\| \leq M_i$, then $\|f^{(i)}\|_{L^\infty} \leq M_i$ for all $i \in \overline{1,m}$.

2.1.3 Hamilton–Cayley theorem in linear algebra

Assume V is a finite dimensional vector space over K ($\mathbb{R}$ or $\mathbb{C}$), for $f \in End_K(V)$, we set

$$p_f(\lambda) = \det(\lambda Id_K - f)$$

and call it is the characteristic polynomial of $f \in End_K(V)$.

Also note that when $\dim_V < \infty$, then $A = [f]_{\mathcal{B}}$ is a $n \times n$ matrix in $M_n(K)$ if $\mathcal{B}$ is any basic of V. So another way to define p_f is

$$p_f(\lambda) = p_A(\lambda) = \det(\lambda I_n - A)$$

and clearly this is a polynomial of degree less or equal than n.

Theorem 2.6 (Hamilton–Cayley)**:** For $f \in End_K(V)$, then $p_f(f) = 0$, in the sense of

$$p(f) = a_n f^n + \ldots + a_1 f + a_0 Id_K$$

if $p(x) = a_n x^n + \ldots + a_1 x + a_0$, and $f^n = f \circ f \circ \ldots \circ f$ (n–times).

2.2 Non-smooth and non-convex analysis

2.2.1 Lipschitz functions

We recall the notion (Frechét) differential of f, given by

Definition 2.7: Assume $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ and $x \in \Omega$, then f is called (Frechét) differentiable at x if there exists a linear mapping $Df(x) : \mathbb{R}^n \longrightarrow \mathbb{R}$ and a real - valued function defined when $\|h\|$ is small such that

$$f(x+h) = f(x) + \langle Df(x), h\rangle + \|h\|\varepsilon(h)$$

where $\varepsilon(h) \longrightarrow 0$ as $h \longrightarrow 0$. Note that we also use $Df(x)$ to denote the represent vector of $Df(x)$ in $\mathbb{R}^n$.

Definition 2.8: A function $f : (E, \|.\|) \longrightarrow (F, \|.\|)$ is called

- Lipschitz at $x \in E$, if there exists $r > 0$ and $M = M(r)$ such that

$$\|f(x) - f(y)\| \leq M\|x - y\|$$

for all $y \in B(x, r)$.

- (Globally) Lipschitz on E if there exists a constant $C > 0$ such that

$$\|f(x) - f(y)\| \leq C\|x - y\|$$

for all $x, y \in E$.

- Locally Lipschitz on E if for every $x \in E$, there exists $r > 0$ and $M = M(x, r)$ such that

$$\|f(x) - f(y)\| \leq M\|x - y\|$$

for all $y \in B(x, r)$.

It's well-known that Locally Lipschitz property implies the continuity of f. One can see that if f is (Fréchet) differentiable at x, then f is Lipschitz at x. Furthermore, we have the Rademacher's theorem

Theorem 2.9 (H.Rademacher's theorem)**:** Assume $\Omega \subset \mathbb{R}^n$ is open. Then any locally Lipschitz function $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ is a.e differentiable in Ω.

Definition 2.10: A function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is non-Lipschitz at x if there exist two sequences $\{x_i\}, \{y_i\}$ such that $x_i \neq y_i$ for all i with the same limit x as $i \longrightarrow \infty$, such that

$$\limsup_{i \longrightarrow \infty} \frac{|f(y_i) - f(x_i)|}{\|y_i - x_i\|} = +\infty$$

Denote the set of all non-Lipschitz points satisfy the above definition by $\mathcal{S}$. We say that f is strictly continuous at a point x if $x \in \mathcal{R}^n \backslash \mathcal{S}$.

Proposition 2.11: Consider $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, we always have the set $\mathcal{S}$ of all non–Lipschitz points of f is closed.

Proof. Let $\{x_k\}_{k=1}^{\infty} \subset \mathcal{S}$, and $x_k \longrightarrow x$ in $\mathbb{R}^n$. For each $k \in \mathbb{N}$, there exists $a_m^k \longrightarrow x_k$ and $b_m^k \longrightarrow x_k$ as $m \longrightarrow \infty$, such that $a_m^k \neq b_m^k$ for all $m \in \mathbb{N}$, and

$$\limsup_{m \longrightarrow \infty} \frac{|f(a_m^k) - f(b_m^k)|}{\|a_m^k - b_m^k\|} = +\infty$$

Thus, we can assume that there exists subsequences $(a_{m_i}^k) \subset (a_m^k)$ and $(b_{m_i}^k) \subset (b_{m_i}^k)$ such that

$$\lim_{i \longrightarrow \infty} \frac{|f(a_{m_i}^k) - f(b_{m_i}^k)|}{\|a_{m_i}^k - b_{m_i}^k\|} = +\infty$$

Therefore exists $n(k) \in \mathbb{N}$ such that

$$\frac{|f(a^k_{m_i}) - f(b^k_{m_i})|}{\|a^k_{m_i} - b^k_{m_i}\|} > k \qquad \forall\, i \geq n(k) \tag{2.1}$$

Since $a^k_{m_i} \longrightarrow x_k$ as $i \longrightarrow \infty$, we can choose a subsequence $a^k_{m_{i_t}}$ such that

$$\left\| a^k_{m_{i_t}} - x_k \right\| \leq \frac{1}{2^t} \qquad \forall\, t \in \mathbb{N}$$

Similarly, we can assume (without loss of generality, just for convenience) $b^k_{m_i} \longrightarrow x_k$ as $i \longrightarrow \infty$, we can choose a subsequence $b^k_{m_{i_t}}$ such that

$$\left\| b^k_{m_{i_t}} - x_k \right\| \leq \frac{1}{2^t} \qquad \forall\, t \in \mathbb{N}$$

One can see that at $t = k + n(k)$, we have

$$\left\| a^k_{m_{i_{k+n(k)}}} - x \right\| \leq \left\| a^k_{m_{i_{k+n(k)}}} - x_k \right\| + \|x_k - x\| \leq \frac{1}{2^{k+n(k)}} + \|x_k - x\| \qquad \forall\, k \in \mathbb{N}$$

Letting $k \longrightarrow \infty$, we can see that

$$\alpha_k = a^k_{m_{i_{k+n(k)}}} \longrightarrow x \qquad \text{as} \qquad k \longrightarrow \infty$$

Doing similarly for $b^k_{m_{i_{k+n(k)}}}$, we have

$$\beta_k = b^k_{m_{i_{k+n(k)}}} \longrightarrow x \qquad \text{as} \qquad k \longrightarrow \infty$$

Clearly $\alpha_k \neq \beta_k$ for all $k \in \mathbb{N}$, furthermore from (2.1) we conclude that

$$\frac{|f(\alpha_k) - f(\beta_k)|}{\|\alpha_k - \beta_k\|} > k \qquad \forall\, k \in \mathbb{N}$$

And thus, we can deduce $x \in \mathcal{S}$, so $\mathcal{S}$ is closed. $\square$

2.2.2 Generalized differentials

Now, we extend the notion (Fréchet) differential into one-sided differential.

Definition 2.12: Let f be a real valued function defined on the open set $\Omega \subset \mathbb{R}^n$. For any $x \in \Omega$, the sets,

$$D^- f(x) = \left\{ p \in \mathbb{R}^n : \liminf_{y \longrightarrow x} \frac{f(y) - f(x) - \langle p, y - x \rangle}{\|y - x\|} \geq 0 \right\}$$
$$D^+ f(x) = \left\{ p \in \mathbb{R}^n : \limsup_{y \longrightarrow x} \frac{f(y) - f(x) - \langle p, y - x \rangle}{\|y - x\|} \leq 0 \right\}$$

are called, respectively the (Frechét) **subdifferential** and **superdifferential** of f at x.

Now we have some basic properties of superdifferential and supdifferential of f

Proposition 2.13: Let $f : \Omega \longrightarrow \mathbb{R}$ and $x \in \Omega$, then the following properties hold

(a) $D^+f(x) = -D^-(-f)(x)$

(b) $D^+f(x)$ and $D^-f(x)$ are convex (possibly empty)

(c) If $f \in C(\Omega)$, then $p \in D^+f(x)$ if and only if there is a function $\varphi \in C^1(\Omega)$ such that $\nabla\varphi(x) = p$ and $f - \varphi$ has a local maximum at x.

(d) If $f \in C(\Omega)$, then $p \in D^-f(x)$ if and only if there is a function $\varphi \in C^1(\Omega)$ such that $\nabla\varphi(x) = p$ and $f - \varphi$ has a local minimum at x.

(e) $D^+f(x)$ and $D^-f(x)$ are both nonempty if and only if f is differentiable at x. In this case we have that

$$D^+f(x) = D^-f(x) = \{\nabla f(x)\}$$

(f) If $f \in C(\Omega)$, the sets of points where a one-sided differential exists

$$\Omega^+ = \{x \in \Omega : D^+f(x) \neq \emptyset\} \qquad \Omega^- = \{x \in \Omega : D^-f(x) \neq \emptyset\}$$

are both non-empty. Indeed, they are dense in Ω.

Proof.

(a) It's clearly since if $a_n \longrightarrow a$ then

$$\limsup_{n\longrightarrow\infty}(-a_n) = -\liminf_{n\longrightarrow\infty} a_n$$

(b) It's also clearly from the definitions.

(c) Assume that $p \in D^+f(x)$, by definition, we can find $\delta > 0$ and a continuous increasing function $\sigma : [0,\infty) \longrightarrow \mathbb{R}$ with $\sigma(0) = 0$ such that

$$f(y) \leq f(x) + \langle p, y - x\rangle + \|y - x\|\sigma(\|y - x\|) \tag{2.2}$$

for $\|y - x\| < \delta$. Define

$$\rho(r) = \int_0^r \sigma(t)\,dt \qquad \Longrightarrow \qquad \rho(0) = \rho'(0) = 0 \quad \text{and} \quad r\sigma(r) \leq \rho(2r) \leq r\sigma(2r)$$

Now for $y \in B(x,\delta)$ we setting

$$\varphi(y) = f(x) + \langle p, y - x\rangle + \rho(2\|y - x\|)$$

Since $f(x) = \varphi(x)$, the fact that φ is differentiable is clearly and since

$$\sigma(r) \leq \frac{\rho(2\|y - x\|)}{\|y - x\|} \leq \sigma(2\|y - x\|)$$

and $\sigma(r) \longrightarrow 0$ as $r \longrightarrow 0$, we conclude that $\nabla\varphi(x) = p$. Now for all $y \in B(x,\delta)$, from (2.2) we have

$$f(y)-f(x) \le \langle p, y-x\rangle + \|y-x\|\sigma(\|y-x\|) \le \langle p, y-x\rangle + \rho(2\|y-x\|) = \varphi(y)-\varphi(x)$$

Therefore, $(f-\varphi)(y) \le (f-\varphi)(x)$ for $y \in B(x,\delta)$, i.e $f-\varphi$ has a local maximum at x.

For the converse, if $\varphi \in C^1(\Omega)$ such that $f-\varphi$ has a local maximum at x and $f(x) = \varphi(x)$, $\nabla\varphi(x) = p$ then since $f(y)-f(x) \le \varphi(y)-\varphi(x)$ in a neighborhood of x, we have

$$\limsup_{y\longrightarrow x} \frac{f(y)-f(x)-\langle p, y-x\rangle}{\|y-x\|} \le \limsup_{y\longrightarrow x} \frac{\varphi(y)-\varphi(x)-\langle p, y-x\rangle}{\|y-x\|} = 0$$

Therefore $p \in D^+f(x)$.

(d) This completely similar to (c).

(e) If f is differentiable at x, then clearly $\nabla f(x) \in D^+f(x) \cap D^-f(x)$. Furthermore, if $p \in D^+f(x)$, then there exists $\varphi \in C^1(\Omega)$ such that

$$\varphi(x) = f(x) \qquad\qquad \nabla\varphi(x) = p$$

and $f-\varphi$ has a local maximum at x, hence $\nabla(f-\varphi)(x) = 0$, therefore $p = \nabla\varphi(x) = \nabla f(x)$. Doing similarly for $D^-f(x)$ we have $D^+f(x) = D^-f(x) = \{\nabla f(x)\}$.

For the converse, assume that $D^+f(x)$ and $D^-f(x)$ are both nonempty. Assume $a \in D^+f(x)$ and $b \in D^-f(x)$, then there exists $\varphi, \psi \in C^1(\Omega)$ such that

$$\varphi(x) = \psi(x) = f(x) \quad \text{and} \quad \begin{cases} \nabla\varphi(x) &= p \qquad f-\varphi \text{ has local maximum at } x \\ \nabla\psi(x) &= q \qquad f-\psi \text{ has local minimum at } x \end{cases}$$

Therefore, in neighborhood $B(x,\delta)$ we have

$$\psi(y) \le f(y) \le \varphi(y) \qquad \forall\, y \quad \text{s.t} \quad \|y-x\| < \delta$$

Since $\psi, \varphi \in C^1(\Omega)$, it's easy to see that f is also differentiable at x, and so we also have the formula $D^+f(x) = D^-f(x) = \{\nabla f(x)\}$.

(f) Let $x_0 \in \Omega$ and $\varepsilon > 0$ be given. We will show that there exists a function $\varphi \in C^1(\Omega)$ such that $f-\varphi$ has local maximum in $B(x_0,\varepsilon)$ at some point y in $B(x_0,\varepsilon)$. Consider the smooth function in $C^1(B(x_0,\varepsilon))$ given by

$$\varphi(x) = \frac{1}{\varepsilon^2 - \|x-x_0\|^2} \qquad x \in B(x_0,\varepsilon)$$

It's easy to extend φ into a function in $C^1(\Omega)$. Also observe that

$$\varphi(x) \longrightarrow +\infty \qquad \text{as} \qquad \|x-x_0\| \longrightarrow \varepsilon-$$

Since f is continuous, we have $f-\varphi$ has a local maximum in $B(x_0,\varepsilon)$, denoted by y. By (c), we conclude that $p = \nabla\varphi(y) \in D^+f(x)$ and therefore $D^+f(x) \neq \emptyset$, i.e $y \in \Omega^+$.

Furthermore, for every $x_0 \in \Omega$ and $\varepsilon > 0$ so small enough, the set Ω^+ contains a point $y \in B(x_0, \varepsilon)$. It's show that Ω^+ is dense in Ω.

Similarly, if we consider the $C^1(B(x_0, \varepsilon))$ function given by

$$\varphi(x) = \frac{-1}{\varepsilon^2 - \|x - x_0\|^2} \qquad x \in B(x_0, \varepsilon)$$

then the case Ω^- is entirely following.

□

If we replace the function φ with $\overline{\varphi}(y) = \varphi(y) \pm \|y - x\|^2$, then (c),(d) in above proposition become $f - \varphi$ has a **strict** local maximum (or **strict** local minimum) instead of local maximum (or local minimum). In view of this case, we have

Lemma 2.14: Let $f : \Omega \longrightarrow \mathbb{R}$ be continuous. Assume for some $\varphi \in C^1(\Omega)$, $f - \varphi$ has a strict local minimum (or a strict local maximum) at $x \in \Omega$. If $f_n \longrightarrow f$ uniformly, then there exists a sequence of points $x_n \longrightarrow x$ with $f_n(x_n) \longrightarrow f(x)$ and such that

$$f_n - \varphi \qquad \text{has a local minimum (or a local maximum) at } x_n$$

for all $n \in \mathbb{N}$.

Proof. Assume $f - \varphi$ has a strict local minimum at x, then for $\delta > 0$ so small enough, there exists $\varepsilon_\delta > 0$ such that

$$f(y) - \varphi(y) > f(x) - \varphi(x) + \varepsilon_\delta \qquad \text{for all } y \in S(x, \delta)$$

where $S(x, \delta) = \{y \in \Omega : \|y - x\| = r\}$. Since $f_n \longrightarrow f$ uniformly in $C(\Omega)$, there exists $n(\delta) \in \mathbb{N}$ such that for all $n \geq n(\delta)$ we have

$$|f_n(y) - f(y)| < \frac{\varepsilon_\delta}{2}$$

for any $y \in S(x, \delta)$. So for $n \geq n(\delta)$ we have

$$f_n(y) - \varphi(y) > f(y) - \varphi(y) - \frac{\varepsilon_\delta}{2} > f(x) - \varphi(x) + \frac{\varepsilon_\delta}{2}$$

for $y \in S(x, \delta)$. It implies that $f_n - \varphi$ has a strict local minimum at some point $x_n \in B(x, \delta)$. Now letting $\delta \longrightarrow 0$ and $\varepsilon_\delta \longrightarrow 0$ we can construct the sequence $\{x_n\}$. □

2.2.3 Convex sets

A subset $A \subset E$ is called **convex** if for any $x, y \in A$, $\lambda x + (1 - \lambda) y \in A$ for all $\lambda \in (0, 1)$. It's also called **strictly convex** if for any $x, y \in A$, $\lambda x + (1 - \lambda) y \in \operatorname{int}(A)$ for all $\lambda \in (0, 1)$. An affine **hyper plane** is a subset H of the form

$$H = [f = \alpha] = \{x \in E : f(x) = \alpha)\}$$

where f is linear from $E \longrightarrow \mathbb{R}$. Sometimes, we denote $\langle f, x \rangle$ instead of $f(x)$. An well–known result about hyperplane states that

Theorem 2.15: The hyperplane $H = [f = \alpha]$ is closed iff f is continuous.

Let A, B be two subsets of E, we say that the hyperplane $H = [f = \alpha]$ **separates** A and B if

$$f(x) \leq \alpha \quad \forall\, x \in A \qquad \text{and} \qquad f(x) \geq \alpha \quad \forall\, x \in B$$

We say f is **strictly separates** A and B if there exists some $\varepsilon > 0$ such that

$$f(x) \leq \alpha - \varepsilon \quad \forall\, x \in A \quad \text{and} \quad f(x) \geq \alpha + \varepsilon \quad \forall\, x \in B$$

Now since the dual space of $\mathbb{R}^n$ is also (isometric to) $\mathbb{R}^n$, so from theorem 2.2 we have that for any hyperplane in a sub–vector space E of $\mathbb{R}^n$, there exists a unique vector $\xi \in E$ s.t

$$H = [f = \alpha] = \{x \in E : \langle \xi, x \rangle = \alpha\}$$

In this case, if $\|\xi\| = 1$, we call ξ is the **normal vector** of the hyperplane H.

We have a well–know result about convex set in normed spaces.

Theorem 2.16 (Hahn–Banach theorem - geometric form)**:**

(a) Let A, B be two nomempty convex subsets such that $A \cap B = \emptyset$, assume that one of them is open, then there exists a closed hyperplane that separates A and B.

(b) Let A, B be two nomempty convex subsets such that $A \cap B = \emptyset$, assume that A is closed and B is compact, there exists a closed hyperplane that strictly separates A and B.

(c) [Hahn–Banach for finite–dimensions]. In case E be a sub–vector space of $\mathbb{R}^n$, let A, B be two nomempty convex subsets in E such that $A \cap B = \emptyset$, then there exists a closed hyperplane $H = [f = \alpha]$ separates A and B, i.e there exists a unique vector $\xi \in E$ s.t

$$\langle \xi, x \rangle \leq \alpha \quad \forall\, x \in A \quad \text{and} \quad \langle \xi, y \rangle \geq \alpha \quad \forall\, y \in B$$

An application of this result, yells the existence of **support hyperplane**. We call the set $H = [f = \alpha]$ which is presented by normal vector ξ, is a support hyperplane of point $x_0 \in \partial\Omega$, iff

$$f(x) = \langle \xi, x \rangle \leq f(x_0) = \langle \xi, x_0 \rangle \qquad \forall\, x \in \Omega$$

Theorem 2.17: [Support hyperplane in $\mathbb{R}^n$] Assume E is a sub–vector space of $\mathbb{R}^n$, let $\Omega \subset E$ be a closed, convex set with boundary $\partial\Omega$, then each point $x_0 \in \partial\Omega$ admits a support hyperplane where it's normal vector is in E and isn't equal to 0.

Proof. Let $\{x_n\} \subset E \backslash \Omega$ s.t $x_n \longrightarrow x_0$ as $n \longrightarrow \infty$, then for each x_n we have $\{x_n\}$ is compact, convex and Ω is closed, convex, so by Hahn–Banach theorem, there exists a unique vector $\xi_n \in E$ strictly separate $\{x_n\}$ and Ω, i.e

$$\langle \xi_n, x \rangle < \langle \xi_n, x_n \rangle \qquad \forall x \in \Omega$$

Clearly $\xi_n \neq 0$, so we can choose ξ_n has $\|\xi_n\| = 1$. Since $\{x \in E : \|x\| = 1\}$ is compact in $\mathbb{R}^n$, so it's also compact in E, and there exists a subsequence $\xi_{n_k} \longrightarrow \xi$ in E where $\|\xi\| = 1$. Taking $k \longrightarrow \infty$ we have

$$\langle x_{n_k}, x \rangle < \langle \xi_{n_k}, x_{n_k} \rangle \Longrightarrow \langle \xi, x \rangle \leq \langle \xi, x_0 \rangle$$

for all $x \in \Omega$. This complete the proof since $\|\xi\| = 1$. $\square$

We have a nice geometry result in Hilbert space for closed, convex sets.

Theorem 2.18 (The perpendicular projection)**:** If K is a closed convex subset of a Hilbert space H, then for any $x \in H$, there exists a unique $x' \in K$ such that $\|x - x'\| = \min\{\|x - y\| : y \in K\}$,

Proof. Let $\delta = \inf\{\|x - y\| : y \in K\}$ then exists a sequence $x_n \subset K$ such that $\|x - x_n\| \longrightarrow \delta$ as $n \longrightarrow \infty$, using the Parallelogram Law we have

$$\|x_n - x_m\|^2 + 4\left\|x - \frac{x_n + x_m}{2}\right\|^2 = 2\|x - x_m\|^2 + 2\|x - x_m\|^2$$

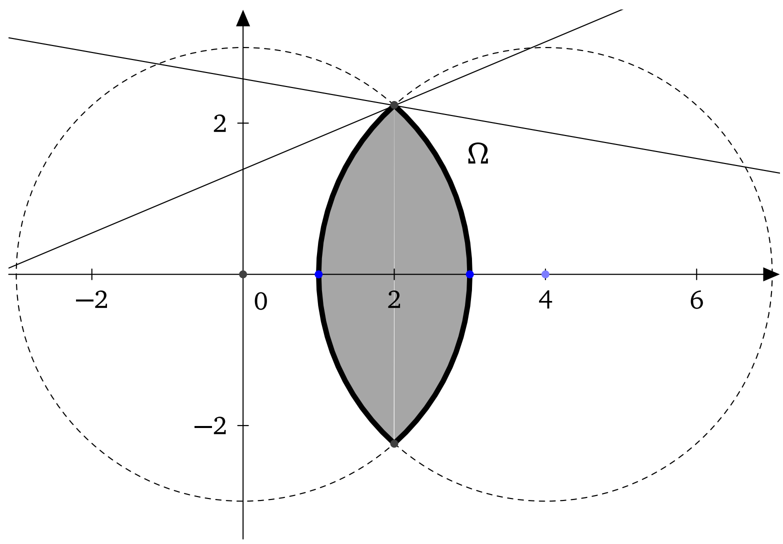
Since K is convex, we have $\frac{x_n+x_m}{2} \in K$, so $\|x_n - x_m\|^2 \leq 2\|x - x_m\|^2 + 2\|x - x_m\|^2 - 4\delta^2$. Let $m, n \longrightarrow \infty$ we conclude that $\{x_n\}$ is a Cauchy sequence, so it's admit a limit point, denoted by x' and $\|x - x'\| = \delta$. Now assume that there exist $a, b \in K$ such that $\|x - a\| = \|x - b\| = \delta$, then from the Parallelogram Law and convexity of K we have

$$\|a - b\|^2 + 4\left\|x - \frac{a+b}{2}\right\|^2 = 2\|x - a\|^2 + 2\|x - b\|^2 \Longrightarrow \|a - b\|^2 \leq 4\delta^2 - 4\left\|x - \frac{a+b}{2}\right\|^2 \leq 0$$

So $a = b$ and the proof is complete. $\square$

We call $x' = \pi(x)$ is the **perpendicular projection** of x to K. Observe that if $0 \notin K$ and choosing $x = 0$, we have $\pi(0)$ is the element which is has minimum norm in K.

Note that if K is convex, closed in E–sub–vector space of $\mathbb{R}^n$, theorem 2.17 claim that at $x_0 \in \partial K$, there exists a support hyperplane $H = [f = \alpha]$ which respect to the normal vector ξ. And easily seen that the support hyperplane is not unique, for example



From this example, we can see that in general, indeed if K is strictly convex, the supporting hyperplane at points in ∂K may be not unique.

But in case K is strictly convex, we have a special property, that is every supporting hyperplane $H = [f = \delta]$ at $x_0 \in \partial K$ such that $H \cap K = \{x_0\}$.

Proposition 2.19: If K is convex, closed in $\mathbb{R}^n$, then K is strictly convex if and only if at any $x_0 \in \partial K$ every supporting hyperplane of K at x_0 has the common with K is exactly $\{x_0\}$.

Proof. First if K is strictly convex and $H = [f = \delta]$ where $\|f\| = 1$ is a supporting hyperplane of K at x_0. Assume by contradiction that there exists another point $y \neq x$ such that $y \in H \cap K$, then by the strictly convexity we have

$$\lambda x + (1-\lambda)y \in \operatorname{int} K \qquad \forall\, \lambda \in (0,1)$$

Clearly, since $f(x) = f(y) = \delta$, we have $f(\lambda x + (1-\lambda)y) = \delta$ for all $\lambda \in (0,1)$, so $\lambda x + (1-\lambda)y \in H \cap \operatorname{int} K$ for all $\lambda \in (0,1)$. Now let $z = \lambda x + (1-\lambda)y$ for fixed $\lambda \in (0,1)$, there exists $r > 0$ such that $B(z,r) \subset K$, and since $f(w) \leq f(x) = \delta$ for all $w \in K$, we have

$$f(z+ra) = f(z) + rf(a) \leq f(x) \quad \forall\, a \in B(0,1) \Longrightarrow f(a) \leq 0 \quad \forall\, a \in B(0,1) \Longrightarrow f \equiv 0$$

It's yell a contradiction since $\|f\| = 1$. So we must have $H = [f = \delta] \cap K = \{x_0\}$.

Conversely, if for any $x_0 \in \partial K$, every supporting hyperplane of K at x_0 have the intersection with K exactly $\{x_0\}$, we will prove that K is strictly convex. Indeed, Assume there exist $x, y \in K$ such that there exists $\lambda \in (0,1)$ such that $z = \lambda x + (1-\lambda)y \in \partial K$. There exists a supporting hyperplane which is respect to the unit normal vector ξ, and using the assumption we have

$$\langle \xi, w\rangle < \langle \xi, z\rangle \qquad \forall\, w \in K\backslash\{z\}$$

So $\langle \xi, x\rangle < \langle \xi, z\rangle$ and $\langle \xi, y\rangle < \langle \xi, z\rangle$, so clearly

$$\langle \xi, z\rangle = \langle \xi, \lambda x + (1-\lambda)y\rangle < \lambda\langle \xi, x\rangle + (1-\lambda)\langle \xi, y\rangle = \langle \xi, z\rangle$$

It's a contradiction, so K must be strictly convex. □

2.2.4 Proximal vectors of closed sets

Definition 2.20: In case $K \subset \mathbb{R}^n$ is closed and convex with boundary ∂K, the normal cone of K at x is defined by

$$N_K(x) = \{w \in \mathbb{R}^n : \langle w, y - x\rangle \leq 0 \qquad \forall\, y \in K\}$$

For a closed set K in $\mathbb{R}^n$, it's well define the distance function

$$d_K(x) = \min_{s \in K} \|x - s\|$$

Also define the projection

$$\operatorname{proj}_K(x) = \{y \in K : \|y - x\| = d_K(x)\}$$

Definition 2.21: For $K \subset \mathbb{R}^n$ is closed with boundary ∂K, for every $x \in K$, we call $\zeta \in \mathbb{R}^n$ is a proximal normal vector of K at x if there exists $r > 0$ such that

$$d_K(x + r\zeta) = r\|\zeta\|$$

and define the set of proximal normal vector ζ is the proximal cone

$$N_K^P(x) = \{\zeta \in \mathbb{R}^n : \zeta \text{ is a proximal vector of } K \text{ at } x\}$$

It's well-known that N_K^P is convex. The next proposition give us another equivalent definitions of $N_k^P(x)$, which is very useful later.

Proposition 2.22: For $K \subset \mathbb{R}^n$ is closed with boundary ∂K, for $x \in K$ and $0 \neq \zeta \in \mathbb{R}^n$, the followings are equivalent

(i) (Characteristic) $\zeta \in N_K^P(x)$.

(ii) (Globally) There exists $\sigma = \sigma(\zeta, x) \geq 0$ such that

$$\langle \zeta, y - x \rangle \leq \sigma \|y - x\|^2 \qquad \forall y \in K$$

(iii) (Locally) For any $\eta > 0$, there exists $\sigma = \sigma(\zeta, x) \geq 0$ such that

$$\langle \zeta, y - x \rangle \leq \sigma \|y - x\|^2 \qquad \forall y \in K \cap B(x, \eta)$$

Proof. First we prove (i) and (ii) are equivalent.

$$\begin{aligned}
\zeta \in N_K^P(x) &\Longleftrightarrow \exists\, r > 0 : d_K(x + r\zeta) = r\|\zeta\| \\
&\Longleftrightarrow \exists\, r > 0 : \|x + r\zeta - x\| \leq \|x + r\zeta - y\| \quad \forall\, y \in K \\
&\Longleftrightarrow \exists\, r > 0 : \|x + r\zeta - x\|^2 \leq \|x + r\zeta - y\|^2 \quad \forall\, y \in K \\
&\Longleftrightarrow \exists\, r > 0 : -2\langle x + r\zeta, x \rangle + \|x\|^2 \leq -2\langle x + r\zeta, y \rangle + \|y\|^2 \quad \forall\, y \in K \\
&\Longleftrightarrow \exists\, r > 0 : 2r\langle \zeta, y - x \rangle \leq \|x\|^2 + \|y\|^2 - 2\langle x, y \rangle \quad \forall\, y \in K \\
&\Longleftrightarrow \exists\, r > 0 : \langle \zeta, y - x \rangle \leq \frac{1}{2r}\|y - x\|^2 \quad \forall\, y \in K
\end{aligned}$$

The rest is prove (iii) implies (ii). Assume there exists $\sigma > 0$ and $\eta > 0$ such that

$$\langle \zeta, y - x \rangle \leq \sigma \|y - x\|^2 \qquad \forall y \in K \cap B(x, \eta)$$

Pick $y \in K \backslash B(x, \eta)$, there are two cases

- If $\|y - x\| \geq 1$, then clearly

$$\langle \zeta, y - x \rangle \leq \|\zeta\| \cdot \|y - x\| \leq \|\zeta\| \cdot \|y - x\|^2$$

- If $\|y - x\| < 1$, then $\eta \leq \|y - x\| < 1$, so $\frac{\|y-x\|}{\eta} \geq 1$ and $\frac{1}{\eta} > 1$. So

$$\langle \zeta, y - x \rangle \leq \|\zeta\| \cdot \|y - x\| \leq \frac{\|\zeta\|}{\eta} \cdot \|y - x\|^2$$

Now taking $\overline{\sigma} = \max\left\{\sigma, \|\zeta\|, \frac{\|\zeta\|}{\eta}\right\}$ then we have

$$\langle \zeta, y - x \rangle \leq \overline{\sigma} \|y - x\|^2 \qquad \forall\, y \in K$$

The implication (ii) to (iii) is trivial, so the proof is complete. $\square$

In case $\zeta \in N_K^P(x)$, by definition, the set $\{\lambda > 0 : d_K(x+\lambda\zeta) = \lambda\|\zeta\|\}$ is not empty. It's equivalent to

$$\exists\, \lambda > 0 : \quad d_K\left(x + \lambda\|\zeta\|\frac{\zeta}{\|\zeta\|}\right) = \lambda\|\zeta\| \quad \Longleftrightarrow \quad \exists\, \lambda > 0 : \quad K \cap B\left(x + \lambda\|\zeta\|\frac{\zeta}{\|\zeta\|}, r\|\zeta\|\right) = \emptyset$$

Definition 2.23: For $x \in K$ and $0 \neq \zeta \in N_K^P(x)$, we say that

(i) ζ is **realized by a ball of radius** $r > 0$ if

$$K \cap B\left(x + r\frac{\zeta}{\|\zeta\|}, r\right) = \emptyset \qquad \Longleftrightarrow \qquad d_K\left(x + r\frac{\zeta}{\|\zeta\|}\right) = r$$

In view of globally definition in proposition 2.22, it's equivalent to

$$\langle \zeta, y - x\rangle \le \frac{\|\zeta\|}{2r} \cdot \|y - x\|^2 \qquad \forall\, y \in K \tag{2.3}$$

(ii) If (2.3) holds for any $x \in \partial K$ and for all $0 \neq \zeta \in N_K^P(x)$, we say that K is **proximally smooth of radius** r or r-**proximally smooth**.

We easily see that a closed set $K \subset \mathbb{R}^n$ is proximal smooth with radius r if and only if every unit normal vector $\zeta \in N_K^P(x)$ can be realized by a ball of radius r, for every $x \in K$. If $0 \neq \zeta \in N_K^P(x)$ is realized by a ball of radius r, then the best constant σ such that

$$\langle \zeta, y - x\rangle \le \sigma\|y - x\|^2 \qquad \forall\, y \in K$$

is $\frac{\|\zeta\|}{2\rho}$ where

$$\rho = \sup\left\{ r > 0 : d_K\left(x + r\frac{\zeta}{\|\zeta\|}\right) = r\right\}$$

We have another simple observation

Proposition 2.24: In case K is convex and closed, then $N_K^P(x) \equiv N_K(x)$ for any $x \in \partial K$.

Proof. Clearly $N_K(x) \subseteq N_K^P(x)$, assume there exists $\zeta \in N_K^P(x) \backslash N_K(x)$, then there exists $y_0 \in K$ and $\varepsilon > 0$ such that

$$\varepsilon < \langle \zeta, y_0 - x\rangle \le \sigma(\zeta, x)\|y_0 - x\|^2 \Longrightarrow \sigma(\zeta, x) > 0$$

Consider $y_\lambda = \lambda y_0 + (1-\lambda)x$ where $\lambda \in (0,1)$, by the convexity of K, $y_\lambda \in K$ for every $\lambda \in (0,1)$, and by the assumption $\zeta \in N_K^P(x)$, we have

$$\langle \zeta, y_\lambda - x\rangle \le \sigma(\zeta, x)\|y_\lambda - x\|^2 = \lambda^2\sigma(\zeta, x)\|y_0 - x\|^2 \tag{2.4}$$

Also, we have

$$\langle \zeta, y_\lambda - x\rangle = \langle \zeta, \lambda(y_0 - x)\rangle = \lambda\langle \zeta, y_0 - x\rangle > \lambda\varepsilon > 0 \tag{2.5}$$

Now from (2.4) and (2.5) we get

$$\lambda\varepsilon < \lambda^2\sigma(\zeta, x)\|y_0 - x\|^2 \Longrightarrow \varepsilon < \lambda\sigma(\zeta, x)\|y_0 - x\|^2$$

for all $\lambda \in (0,1)$. Letting $\lambda \longrightarrow 0$ we get a contradiction. $\square$

Definition 2.25 (Proximal subdifferential and horizon subdifferential)**:** Let $\Omega \subset \mathbb{R}^n$ be open and $f : \Omega \longrightarrow \mathbb{R} \cup \{+\infty\}$ be lower semi-continuous. We define some notions

(i) The epigraph of f is

$$\text{epi}(f) = \{(x, y) \in \Omega \times \mathbb{R} : f(x) \leq y\}$$

(ii) The hypograph of f is

$$\text{hypo}(f) = \{(x, y) \in \Omega \times \mathbb{R} : f(x) \geq y\}$$

(iii) The notion dom(f) of f is

$$\text{dom}(f) = \{x \in \Omega : f(x) < +\infty\}$$

(iv) The proximal subdifferential $\partial_P f(x)$ of f at a point $x \in \text{dom}(f)$ is the set

$$\partial_P f(x) = \left\{\zeta \in \mathbb{R}^n : (\zeta, -1) \in N^P_{\text{epi}(f)}(x, f(x))\right\}$$

(v) The proximal superdifferential $\partial^P f(x)$ of f at a point $x \in \text{dom}(f)$ is the set

$$\partial^P f(x) = \left\{\zeta \in \mathbb{R}^n : (-\zeta, 1) \in N^P_{\text{hypo}(f)}(x, f(x))\right\}$$

(vi) The horizon subdifferential $\partial_\infty f(x)$ of f at a point $x \in \text{dom}(f)$ is the set

$$\partial_\infty f(x) = \left\{\zeta \in \mathbb{R}^n : (\zeta, 0) \in N^P_{\text{epi}(f)}(x, f(x))\right\}$$

(v) The horizon superdifferential $\partial^\infty f(x)$ of f at a point $x \in \text{dom}(f)$ is the set

$$\partial^\infty f(x) = \left\{\zeta \in \mathbb{R}^n : (-\zeta, 0) \in N^P_{\text{hypo}(f)}(x, f(x))\right\}$$

We will present a nice relation between the proximal subdifferential $\partial_P f(x)$ and the the one-sided subdifferential already presented in definition 2.12. Recall from theorem 1.2.5 in [2], we know that

Theorem 2.26: A vector $v \in \mathbb{R}^n$ is belong to $\partial_P f(x)$ at a point $x \in \text{dom}(f)$ if and only if there exist positive numbers σ and η such that

$$f(y) - f(x) - \langle v, y - x\rangle \geq -\sigma \|y - x\|^2 \qquad \forall\ y \in B(x, \eta)$$

In this case, we have

$$\frac{f(y) - f(x) - \langle v, y - x\rangle}{\|y - x\|} \geq -\sigma \|y - x\| \qquad \forall\ y \in B(x, \eta)$$

So letting $y \longrightarrow x$ we have

$$\liminf_{y \longrightarrow x} \frac{f(y) - f(x) - \langle v, y - x\rangle}{\|y - x\|} \geq 0 \qquad \Longrightarrow \qquad v \in D^- f(x)$$

in view of definition 2.12, so we conclude

Proposition 2.27: Consider $f : \mathbb{R}^n \longrightarrow \mathbb{R} \cup \{+\infty\}$, assume $x \in \text{dom}(f)$, then

$$\partial_P f(x) \subset D^- f(x)$$

Now we will characterize some properties of $N^P_{\text{epi}(f)}(x, r)$.

Proposition 2.28: Let $f : \Omega \longrightarrow (-\infty, +\infty]$ is continuous, where $\Omega \subset \mathbb{R}^n$ is open. Suppose that

$$(\zeta, -\lambda) \in N^P_{\text{epi}(f)}(x, r)$$

where $(x, r) \in \text{epi}(f)$, and $\zeta \in \mathbb{R}^n$. Show that $\lambda \geq 0$. Furthermore, If $r > f(x)$ then $\lambda = 0$. Thus If $\lambda > 0$ then $r = f(x)$.

Proof. By definition, $(\zeta, -\lambda) \in N^P_{\text{epi}(f)}(x, r)$ iff there exists $t > 0$ such that

$$d_{\text{epi}(f)}((x, r) + t(\zeta, -\lambda)) = t\|(\zeta, -\lambda)\|$$

It's equivalent to

$$\|(x + t\zeta, r - t\lambda) - (y, \alpha)\| \geq t\|(\zeta, -\lambda)\|$$

for all $f(y) \leq \alpha$. If $\lambda < 0$, then $-t\lambda > 0$, so $\alpha = r - t\lambda > r \geq f(x)$, thus $(x, r - t\lambda) \in \text{epi}(f)$, hence

$$\|(x + t\zeta, r - t\lambda) - (x, r - t\lambda)\| \geq t\|(\zeta, -\lambda)\| \Longrightarrow \|(t\zeta, 0)\| \geq t\|(\zeta, -\lambda)\|$$

It's a contradiction, so we must have $\lambda \geq 0$. If $r > f(x)$, and $\lambda > 0$, take $\alpha = r - \varepsilon > f(x)$ to make sure that $(x, r - \varepsilon) \in \text{epi}(f)$, then

$$\|(x + t\zeta, r - t\lambda) - (x, r - \varepsilon)\| \geq t\|(\zeta, -\lambda)\| \Longrightarrow |\varepsilon - t\lambda| \geq t|\lambda|$$

We can choose ε such that $0 < \varepsilon < \min\{t\lambda, r - f(x)\}$, then it follows a contradiction. And the rest is trivial. $\square$

2.2.5 Convex and concave functions

A function $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ is called convex if and only if

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y) \tag{2.6}$$

for all x, y such that $[x, y] \subset \Omega$ and for all $\lambda \in [0, 1]$.

As a standard criterion of convex functions, we have

Proposition 2.29: A function f is convex if and only if f satisfies (2.6) for $\lambda = \frac{1}{2}$.

Proof. First we prove that if (2.6) is true for $\lambda = \frac{1}{2}$ then it's also true for all $\lambda = \frac{k}{2^n}$ for $1 \leq k \leq 2^n - 1$. Using induction by n

- $n = 2$. We need to prove (2.6) for $\lambda = 1, 2, 3$. We have
  - $k = 1$, $f(x) + 3f(x) = f(x) + f(y) + 2f(y) \leq 2f\left(\frac{x+y}{2}\right) + 2f(y) \leq 4f\left(\frac{x+3y}{4}\right)$

- $k = 2$, this case reduces to $\lambda = \frac{1}{2}$.
- $k = 3$, this case is similar to $k = 1$.

- Assume (2.6) is true for all $1 \leq k \leq 2^n - 1$. For $2^n \leq k \leq 2^{n+1} - 1$ we have
  - If k is even, write $k = 2m$, since $1 \leq m \leq 2^n - 1$ we have

$$2\left[mf(x) + (2^n - m)f(y)\right] \leq 2.2^n f\left(\frac{mx + (2^n - m)y}{2^n}\right) = 2^{n+1} f\left(\frac{2mx + (2^{n+1} - 2m)y}{2^{n+!}}\right)$$

  - If k is odd, since $k \geq 2^n + 1$, clearly $2^{n+1} - k \leq 2^n - 1$, and $k - 2^n = 2^n - (2^n - k)$

$$\begin{aligned} kf(x) + (2^{n+1} - k)f(y) &= \Big((k - 2^n)f(x) + (2^{n+1} - k)f(y)\Big) + 2^n f(x) \\ &\leq 2^n f\left(\frac{(k - 2^n)x + (2^{n+1} - k)y}{2^n}\right) + 2^n f(x) \\ &\leq 2^{n+1} f\left(\frac{kx + (2^{n+1} - k)y}{2^{n+1}}\right) \end{aligned}$$

So (2.6) is true for all $\lambda = \frac{k}{2^n}$, but $\left\{\frac{k}{2^n} :, k, n \in \mathbb{N}\right\}$ is dense in $[0, 1]$, so by continuity, we complete the proof. □

Now we establish some well known facts about convex function

Theorem 2.30: If $\Omega \subset \mathbb{R}^n$ is an open set and $f : \Omega \longrightarrow \mathbb{R}$ is convex, then f is locally bounded in Ω. Furthermore, if $|f(x)| \leq M$ on $B(a, r)$ then f is Lipschitz on $B(a, r - \varepsilon)$ with constant $\frac{2M}{\varepsilon}$. Therefore f is also locally Lipschitz in Ω.

Proof. For any fixed $x_0 \in \Omega$, by the openness, there exists a closed cube K centered at x_0 with vertices $v_1, v_2, \ldots, v_{2^n}$. For any $x \in K$, there exists $\lambda_1, \ldots, \lambda_{2^n} \in [0, 1]$ such that

$$x = \sum_{i=1}^{2^n} \lambda_i v_i \qquad \text{and} \qquad \sum_{i=1}^{2^n} \lambda_i = 1$$

Since f is convex in Ω, we have

$$f(x) \leq \sum_{i=1}^{2^n} \lambda_i f(v_i) \leq \max_{1 \leq i \leq 2^n} f(v_i) = M$$

So f is bounded above in K. On the other hand, for any $x \in K$, by the symmetry of K, there exists $y \in K$ such that $x_0 = \frac{x+y}{2}$, then

$$2f(x_0) \leq f(x) + f(y) \qquad \Longrightarrow \qquad f(x) \geq 2f(x_0) - f(y) \geq 2f(x_0) - M$$

So f is also bounded below in K, hence f is locally bounded in Ω.
For the rest, if $x, y \in B(a, r - \varepsilon)$, by the openness, there exists $\varepsilon > 0$ small enough such that

$z = y + \frac{\varepsilon}{\|y-x\|}(y-x) \in B(a,r)$, so we have

$$\begin{aligned} y = \left(\frac{\varepsilon}{\varepsilon + \|y-x\|}\right)x + \left(\frac{\|y-x\|}{\varepsilon + \|y-x\|}\right)z &\Longrightarrow (\varepsilon + \|y-x\|)f(y) \le \varepsilon f(x) + \|y-x\| f(z) \\ &\Longrightarrow \varepsilon(f(y) - f(x)) \le (f(z) - f(x))\|y-x\| \\ &\Longrightarrow |f(y) - f(x)| \le \frac{2M}{\varepsilon}\|y-x\| \end{aligned}$$

So f is locally Lipschitz on Ω. □

Theorem 2.31: Assume $\Omega \subset \mathbb{R}^n$ is an open set and $f : \Omega \longrightarrow \mathbb{R}$ is a function. Then f is convex if and only if the epigraph of f is convex in $\Omega \times \mathbb{R}$

Proof. Let (a,b) and (x,y) both belong to epi(f), for any $\lambda \in (0,1)$, we have

$$\lambda(a,b) + (1-\lambda)(x,y) = \Big(\lambda a + (1-\lambda)x, \lambda b + (1-\lambda)y\Big)$$

Recall that $f(a) \le b$ and $f(x) \le y$, we have

$$f(\lambda a + (1-\lambda)x) \le \lambda f(a) + (1-\lambda)f(x) \le \lambda b + (1-\lambda)y$$

Therefore epi(f) is also convex. For the converse, since $(a, f(a))$ and $(x, f(x))$ lie in epf(f), which is convex, thus for $\lambda \in (0,1)$ we have

$$\lambda(a, f(a)) + (1-\lambda)(x, f(x)) \in \text{epi}(f) \Longrightarrow f(\lambda a + (1-\lambda)x) \le \lambda f(a) + (1-\lambda)f(x)$$

The proof is complete. □

On the contrary, a function $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ is called concave if and only if

$$f(\lambda x + (1-\lambda)y) \ge \lambda f(x) + (1-\lambda)f(y) \tag{2.7}$$

for all x, y such that $[x,y] \subset \Omega$ and for all $\lambda \in [0,1]$.

In other word, f is convex if and only if $-f$ is concave. So theorem 2.30 are also true for concave function f, by the same argument via convex function $-f$.

Theorem 2.32: If $\Omega \subset \mathbb{R}^n$ is an open set and $f : \Omega \longrightarrow \mathbb{R}$ is concave, then f is locally bounded in Ω. Furthermore, if $|f(x)| \le M$ on $B(a,r)$ then f is Lipschitz on $B(a, r-\varepsilon)$ with constant $\frac{2M}{\varepsilon}$. Therefore f is also locally Lipschitz in Ω.

2.2.6 Semiconvex and semiconcave functions

Now, we extend the notion of concave functions define the notion of semiconcave function. Similarly, f is semiconvex if $-f$ is semiconcave.

Definition 2.33 (Semiconcave functions with linear modulus)**:** Let $\Omega \subset \mathbb{R}^n$ be an open set. We say that a function $f : \Omega \longrightarrow \mathbb{R}$ is semiconcave with linear modulus if it is continuous in Ω and there exists $C > 0$ such that

$$\lambda f(x) + (1-\lambda) f(y) - f(\lambda f(x) + (1-\lambda) y) \leq \lambda(1-\lambda) C \|y - x\|^2 \tag{2.8}$$

for all $\lambda \in [0,1]$ such that $[x,y] \subset \Omega$. The constant C is called a semiconcavity constant for f in Ω.

It's easy to see that when $C = 0$ in (2.8) then f is concave in Ω. Hence a semiconcave function is a concave function up to a quadratic term, more precisely

Lemma 2.34: The function $f : \Omega \longrightarrow \mathbb{R}$ is semiconcave with the semiconcavity constant C in Ω if and only if $f(x) - C\|x\|^2$ is concave in Ω.

Proof. Setting $g(x) = f(x) - C\|x\|^2$, and the rest is trivial by substituting g into (2.8). □

Now we introduce a standard criterion of semiconcave functions

Proposition 2.35: Let $f : \Omega \longrightarrow \mathbb{R}$, then the followings are equivalent

(i) f is continuous and satisfy

$$f(x+h) + f(x-h) - 2f(x) \leq 2C\|h\|^2$$

for all $[x-h, x+h] \subset \Omega$.

(ii) f is semiconcave with a semiconcavity constant C in Ω.

Proof. Setting $g(x) = f(x) - f(x) - C\|x\|^2$, then g is continuous and

$$g(x+h) + g(x-h) \leq 2g(x) \qquad \forall\, [x-h, x+h] \subset \Omega$$

Now using the same argument in proposition 2.29 we can see g is concave, therefore f is semi-concave.
For the converse, if f is semiconcave with semiconcavity constant C, setting $g(x) = f(x) - f(x) - C\|x\|^2$ then g is concave, so by theorem 2.32 we have g is locally Lipschitz in Ω. This implies the continuity of g, and so does f. The rest is trivial, since g is concave, we have

$$g(x+h) + g(x-h) \leq 2g(x) \qquad \forall\, [x-h, x+h] \subset \Omega$$

Combine this fact and $g(x) = f(x) - C\|x\|^2$ we achieve (i). □

We also have the locally Lipschitz of a semiconcave function as follow

Theorem 2.36: If a function $f : \Omega \longrightarrow \mathbb{R}$ is semiconcave with a semiconcavity C in Ω, then f is locally Lipschitz in Ω.

Proof. Setting $g(x) = f(x) - C\|x\|^2$ is Lipschitz with constant M and bounded by m in a neighborhood $K \subset \Omega$, then for $x, y \in K$

$$|g(x) - g(y)| \leq M\|x-y\| \Longrightarrow |f(x) - f(y)| \leq \frac{C}{2}\|x^2 - y^2\| + M\|x-y\| \leq (mC+M)\|x-y\|$$

So f is also locally Lipschitz in Ω. □

Combine with H.Radamacher's theorem 2.9, we have the following result on differentiability of semiconcave function.

Corollary 2.37: A semiconcave function $f : \Omega \longrightarrow \mathbb{R}$ is a.e differentiable in Ω.

For example, we now establish some properties about the distance function from a point to the closed set $S \subset \mathbb{R}^n$, given by

Proposition 2.38: Let $S \subset \mathbb{R}^n$ be a closed set, consider the function

$$d_S(x) = \inf_{y \in S} \|y - x\| \qquad x \in \mathbb{R}^n$$

Prove that

(i) d_S is locally semiconcave in $\mathbb{R}^n \backslash S$.

(ii) d_S is not locally semiconcave in $\mathbb{R}^n$

(iii) d_S^2 is semiconcave with constant 2.

Proof. First we observe that d_S is continuous in $\mathbb{R}^n$. Indeed, for $x, y \in \mathbb{R}^n$, let a be fixed in S, then clearly

$$\|x - a\| \leq \|x - y\| + \|y - a\| \Longrightarrow d_S(x) \leq \|x - a\| \leq \|x - y\| + \|y - a\|$$

From this, for all $a \in S$ we have

$$\|y - a\| \leq \|x - y\| - d_S(x) \Longrightarrow d_S(y) \leq \|y - a\| \leq \|x - y\| - d_S(x)$$

There fore we have

$$d_S(x) - d_S(y) \leq \|x - y\|$$

Doing similarly we easily conclude that d_S is Lipschitz with constant 1, so clearly d_S is continuous in $\mathbb{R}^n$.

(i) If $x_0 \in \mathbb{R}^n \backslash S$, then clearly $d_S(x_0) > 0$ since S is closed. Denote $2r = d_S(x_0) > 0$, for $y \in B(x_0, r)$ and $a \in S$ we have

$$2r = d_S(x_0) \leq \|x_0 - a\| \leq \|x_0 - y\| + \|y - a\| \leq r + \|y - a\| \Longrightarrow \|y - a\| \geq r$$

This is true for any $a \in S$, so $d_S(y) \geq r > 0$ for all $y \in B(x_0, r)$. Now for $x \in B(x_0, r)$ and h such that $[x - h, x + h] \subset B(x_0, r)$, let $a \in S$, we always have $\|x - a\| \geq r$, so

$$\begin{aligned}\Big(\|x + h - a\| + \|x - h - a\|\Big)^2 &\leq 2\Big(\|x + h - a\|^2 + \|x - h - a\|^2\Big) \\ &= 4\Big(\|x - a\|^2 + \|h\|^2\Big) \leq \left(2\|x - a\| + \frac{\|h\|^2}{\|x - a\|}\right)^2\end{aligned}$$

Therefore

$$\|x + h - a\| + \|x - h - a\| \leq 2\|x - a\| + \frac{\|h\|^2}{\|x - a\|} \leq 2\|x - a\| + \frac{\|h\|^2}{r}$$

From proposition 2.35, d_S is semiconcave in $B(x_0, r)$ with semiconcavity $C = \frac{1}{r}$.

(ii) It's easy to see an counter-example, in case $S = \{0\}$ and $x\mathbb{R}^n \backslash 0$, then $d_S(x) = \|x\| > 0$, we will show that d_S is not semiconcave in any neighborhood of 0. Indeed, for any $r > 0$ and $h \in B\left(0, \frac{r}{2}\right)$, if d_S is semiconcave in $B(0,r)$, since $[0-h, 0+h] \subset B(0,r)$, there must be exists $C > 0$ such that

$$d_s(0+h) + d_S(0-h) - 2d_S(0) = 2h \leq Ch^2 \Longrightarrow h \geq \frac{2}{C}$$

Let $h \longrightarrow 0$ we have a contradiction. Therefore d_S is not semiconcave in any neighborhood $B(0,r)$ of 0.

(iii) It's easy to see that

$$d_S^2(x) = \left(\inf_{a \in S} \|x-a\|\right)^2 = \inf_{a \in S} \|x-a\|^2$$

For $x, h \in \mathbb{R}^n$ and $a \in S$ be fixed, then $\|x+h-a\|^2 + \|x-h-a\|^2 - 2\|x-a\|^2 = 2\|h\|^2$

$$d_S^2(x+h) + d_S^2(x-h) \leq \|x+h-a\|^2 + \|x-h-a\|^2 = 2\|h\|^2 + 2\|x-a\|^2$$

Since this equation is true for all $a \in S$, we conclude that

$$d_S^2(x+h) + d_S^2(x-h) \leq 2\|h\|^2 + 2d_S^2(x)$$

and from proposition 2.35, d_S^2 is semiconcave in $\mathbb{R}^n$ with semiconcavity $C = 2$.

□

2.2.7 Sets with positive reach

In this section, we introduce some concepts, due to [13], given an arbitrary closed set $K \subset \mathbb{R}^n$, we set

$$\mathrm{Unp}(K) = \{x \in \mathbb{R}^n : \mathrm{proj}_K(x) \text{ is a singleton}\}$$
$$\mathrm{reach}(K,x) = \sup\left\{r \geq 0 : B(x,r) \subseteq \mathrm{Unp}\ (K)\right\} \qquad \forall\ x \in K$$

and define the reach of K to be

$$\mathrm{reach}(K) = \inf_{x \in K}\ \mathrm{reach}(K,x)$$

Definition 2.39: We say that a closed set $K \subseteq \mathbb{R}^n$ has **positive reach** if $\mathrm{reach}(K) > 0$. Similarly, a closed $K \subseteq \mathbb{R}^n$ has **locally positive reach** if $\mathrm{reach}(K,x) > 0$ for all $x \in K$.

From this definition, we can define the mapping $x \longmapsto \pi_K(x) = \mathrm{proj}_K(x)$ for $x \in \mathrm{Unp}(K)$. Furthermore, for $x \in \partial K$, then

Proposition 2.40: For a closed $K \subseteq \mathbb{R}^n$, we have $\mathrm{Unp}(\partial K) \subseteq \mathrm{Unp}(K)$.

Proof. Assume $x \in \mathrm{Unp}(K)$, if $x \in K$ then clearly $x \in \mathrm{Unp}(K)$. Consider $x \notin K$, then exists unique $a \in \partial K$ such that

$$\mathrm{proj}_{\partial K}(x) = \{a\} \qquad \text{i.e,} \qquad d_{\partial K}(x) = \|x-a\|$$

If $x \notin \mathrm{Unp}(K)$, then there exists an element $y \in \mathrm{proj}_K(x)$ such that $y \neq a$.

- If $y \in \partial K$, then it's a contradiction since $\text{proj}_{\partial K}(x) = \{a\}$
- If $y \in K \backslash \partial K$, then y must lie in the interior of K, so there exists $\varepsilon > 0$ such that $B(y, \varepsilon) \subset K$. Define $r = \|x - y\| = d_K(x) > 0$. We must have

  $$r = \|x - y\| \leq \|x - z\| \qquad \forall\, z \in K$$

  Consider the set $\{y + \zeta : \zeta \in B(0, \varepsilon)\} \subset B(y, \varepsilon) \subset K$, we have

  $$r = \|x - y\| \leq \|x - y - \zeta\| \qquad \forall\, \zeta \in B(0, \varepsilon) \tag{2.9}$$

  Let $\zeta = \frac{\varepsilon(x-y)}{2r}$, then

  $$\|\zeta\| = \frac{\|x - t\|}{r}\frac{\varepsilon}{2} = \frac{\varepsilon}{2} < \varepsilon \Longrightarrow \zeta \in B(0, \varepsilon)$$

  So substitute this into (2.9) we have

  $$r = \|x - y\| \leq \left\| x - y - \frac{\varepsilon(x - y)}{2r} \right\| = \left(1 - \frac{\varepsilon}{2r}\right) \|x - y\| < \|x - y\| = r$$

  It's a contradiction.

In summary, we must have $\text{proj}_K(x) = \{a\} = \text{proj}_{\partial K}(x)$, i.e $\text{Unp}(\partial K) \subseteq \text{Unp}(K)$. $\square$

Thus, for $x \in \partial K$, we deduce that

$$0 \leq \text{reach}(\partial K, x) \leq \text{reach}(K, x) \leq +\infty$$

In case K is closed and convex, clearly every point in $\mathbb{R}^n$ has the unique projection into K, so $\text{Unp}(K) = \mathbb{R}^n$, thus $\text{reach}(K) = +\infty$, since $\text{reach}(K, x) = +\infty$ for any $x \in K$.

The converse is also true, and is the famous result which is known as the Motzkin's theorem

Theorem 2.41 (Motzkin's theorem)**:** Let $K \subset \mathbb{R}^n$ be closed. If for every $x \in \mathbb{R}^n$, the projection $\pi_K(x)$ is unique determined, then K is convex.

From this, if K has locally positive reach, i.e $\text{reach}(A, x) > 0$ for all $x \in K$, then $\text{Unp}(K) = \mathbb{R}^n$, by theorem 2.41 we have K is convex. Thus we obtain

Proposition 2.42: If $K \subset \mathbb{R}^n$ is closed, then K is convex if and only if $\text{reach}(K) = +\infty$.

We have a simple observation

Proposition 2.43: The function

$$\begin{aligned} \text{reach}(K, .) : K &\longrightarrow [0, +\infty] \\ x &\longmapsto \text{reach}(K, x) \end{aligned}$$

is continuous.

Proof. Let $x \in K$, given $\varepsilon > 0$, we claim that for all $y \in B\left(x, \frac{\varepsilon}{2}\right)$ then $|\text{reach}(K, x) - \text{reach}(K, y)| \leq \varepsilon$.

- If $\text{reach}(K,x) = 0$, then for any $r > 0$, we have $B(x,r) \cap \mathbb{R}^n \backslash \text{Unp}(K) \neq \emptyset$. We have

$$y \in B\left(x, \frac{\varepsilon}{2}\right) \Longrightarrow \text{reach}(K,y) \leq \varepsilon$$

Indeed, for any $r > \varepsilon$, then $B\left(x, \frac{\varepsilon}{2}\right) \subseteq B(y,r)$ since

$$z \in B\left(x, \frac{\varepsilon}{2}\right) \Longleftrightarrow \|z-x\| < \frac{\varepsilon}{2} \Longrightarrow \|z-y\| \leq \|z-x\| + \|x-y\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon < r$$

Therefore

$$\emptyset \neq B\left(x, \frac{\varepsilon}{2}\right) \subseteq B(y,r) \cap \mathbb{R}^n \backslash \text{Unp}(K)$$

So we must have $\text{reach}(K,y) \leq \varepsilon$, thus $|\text{reach}(K,x) - \text{reach}(K,y)| \leq \varepsilon$.

- If $\text{reach}(K,x) = r > 0$. We only need to consider $\varepsilon < r$. We have

$$B(y, r-\varepsilon) \subseteq B\left(x, r - \frac{\varepsilon}{2}\right) \subseteq \text{Unp}(K) \Longrightarrow \text{reach}(K,y) \geq r - \varepsilon$$

Furthermore, for any $\eta > r + \varepsilon$, one has $B\left(x, r + \frac{\varepsilon}{2}\right) \subseteq B(y,\eta)$ since

$$z \in B\left(x, r + \frac{\varepsilon}{2}\right) \Longrightarrow \|z-y\| \leq \|z-x\| + \|x-y\| \leq r + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = r + \varepsilon < \eta$$

Thus

$$\emptyset \neq B\left(x, r + \frac{\varepsilon}{2}\right) \cap \mathbb{R}^n \backslash \text{Unp}(K) \subseteq B(y,\eta) \cap \mathbb{R}^n \backslash \text{Unp}(K)$$

So we must have $\text{reach}(K,y) \leq r + \varepsilon$, and thus

$$r - \varepsilon \leq \text{reach}(K,y) \leq r + \varepsilon \Longrightarrow |\text{reach}(K,x) - \text{reach}(K,y)| \leq \varepsilon$$

Thus $x \longmapsto \text{reach}(K,x)$ is continuous in K. □

In case K is closed and has positive reach, we have a better bound $\sigma > 0$ in proximal smooth vector definition, in view of proposition 2.22. This result come from [13], theorem 4.8 (7).

Lemma 2.44: For non-empty closed subset K of $\mathbb{R}^n$, if $x \in \text{Unp}(K)$ has $\pi_K(x) = a$, $\text{reach}(K,a) > 0$, then for any $y \in A$, we have

$$\langle x-a, y-a \rangle \leq \frac{\|x-a\|}{2\,\text{reach}(K,a)} \|y-a\|^2$$

Proposition 2.45: If $K \subset \mathbb{R}^n$ is closed, and has locally positive reach, i.e $\text{reach}(K,x) > 0$ for all $x \in K$, for any $\zeta \in N_K^p(x)$, we have

$$\left\langle \frac{\zeta}{\|\zeta\|}, y-x \right\rangle \leq \frac{1}{2\,\text{reach}(K,x)} \|y-x\|^2 \qquad \forall\, y \in K$$

Proof. Since $0 \neq \zeta \in N_K^p(x)$, there exists $r > 0$ such that $d_K(x + r\zeta) = r\|\zeta\|$, thus

$$x \in \operatorname{proj}_K(x + r\zeta)$$

Reducing $r > 0$ such that $x + r\zeta \in B(x,s) \subset \operatorname{Unp}(K)$ for some $s < textreach(K,x)$, we can assume

$$x + r\zeta \in \operatorname{Unp}(K) \qquad \text{and} \qquad \pi_K(x + r\zeta) = r\|\zeta\|$$

Now using lemma 2.44 we obtain

$$\langle \zeta, y - x \rangle \leq \frac{\|\zeta\|}{2\operatorname{reach}(K,x)}\|y - x\|^2 \qquad \forall\, y \in A$$

Thus the proof is complete. □

Now we state some relation definitions

Definition 2.46: Let $K \subset \mathbb{R}^n$ be closed. Then K is φ-**convex** if and only if there exists a continuous function $\varphi : K \longrightarrow (0, +\infty]$ such that the inequality

$$\langle \zeta, y - x \rangle \leq \varphi(x)\|\zeta\|\|y - x\|^2$$

holds for all $x, y \in K$ and $\zeta \in N_K^P(x)$. In other word, K is φ-convex if and only at every point $x \in \partial K$, every non-zero vector $\zeta \in N_K^P(x)$ can be realized by a ball of radius $\frac{1}{2\varphi(x)}$. In case $\varphi \equiv \varphi_0$ is a constant, we can see that K is φ_0-convex it equivalent to K is $\frac{1}{2\varphi_0}$-proximally smooth.

By proposition 2.45, we can see that if K is non-empty, closed in $\mathbb{R}^n$ and has locally positive reach, then

$$\left\langle \frac{\zeta}{\|\zeta\|}, y - x \right\rangle \leq \frac{1}{2\operatorname{reach}(K,x)}\|y - x\|^2 \qquad \forall\, y \in K$$

holds. Combine with the fact that $x \longmapsto \operatorname{reach}(K,x)$ is continuous in proposition 2.43, we obtain

$$K \text{ has locally positive reach} \implies K \text{ is } \frac{1}{2\operatorname{reach}(K,.)} - \text{convex}$$

In case K has positive reach, then

$$K \text{ has positive reach} \implies K \text{ is } \frac{1}{2\operatorname{reach}(K)} - \text{convex}$$

It' well study in [12] that the converse is also true. And s a closed set K is proximal smooth, or φ_0-convex if and only if K has positive reach. Thus, these definitions (and a various other) are equivalent. These also related with the following weaker definition.

Definition 2.47: Let $K \subseteq \mathbb{R}^n$ be closed and $\theta : \partial K \longrightarrow (0, \infty]$ be continuous. We say that K satisfies the θ-external sphere condition if and only if for every $x \in \partial K$, there exists a non-zero vector $\zeta_x \in N_K^P(x)$ is realized by a ball of radius $\theta(x)$. In other word,

$$\langle \zeta_x, y - x \rangle \leq \frac{\|\zeta_x\|}{2\theta(x)}\|y - x\|^2 \qquad \forall\, y \in K$$

In case $\theta(.) \equiv \rho_0$, we say that K is satisfies the ρ_0-external sphere condition.

We can easily see that

$$K \text{ is } \varphi-\text{convex} \Longrightarrow K \text{ is satisfies the } \frac{1}{2\varphi}-\text{external sphere condition}$$

2.2.8 Semiconvex/concave functions and sub/superdifferentials

Proposition 2.48: Consider $\Omega \subset \mathbb{R}^n$ is an open set. Let $f : \Omega \longrightarrow \mathbb{R}$ is semi-concave with linear modulus C, then a vector $\zeta \in \mathbb{R}^n$ belong to $D^+f(x)$ if and only if

$$f(y) - f(x) - \langle \zeta, y - x \rangle \leq C\|y - x\|^2 \tag{2.10}$$

for any point $y \in \Omega$ such that $[x, y] = \{\lambda x + (1 - \lambda y) : \lambda \in [0, 1]\} \subseteq \Omega$.

Proof. If (2.10) hold, then clearly

$$\limsup_{y \longrightarrow x} \frac{f(y) - f(x) - \langle \zeta, y - x \rangle}{\|y - x\|} \leq 0 \quad \Longrightarrow \quad \zeta \in D^+f(x)$$

For the converse, let $\zeta \in D^+f(x)$, then there exists $\delta > 0$ such that there eixsts a continuous increasing function $\sigma : [0, +\infty) \longrightarrow \mathbb{R}$ with $\sigma(0) = 0$ and

$$f(\overline{y}) - f(x) \leq \langle \zeta, \overline{y} - x \rangle + \|\overline{y} - x\|\sigma(\|\overline{y} - x\|) \qquad \forall\, \overline{y} \in B(x, \delta) \tag{2.11}$$

Recall the semi-concavity property of f, for $[x, y] \subset \Omega$, it holds

$$\lambda f(x) + (1 - \lambda)f(y) \leq f(\lambda x + (1 - \lambda)y) + \lambda(1 - \lambda)C\|y - x\|^2$$

It's equivalent to

$$(1 - \lambda)f(y) - (1 - \lambda)f(x) \leq f(x + (1 - \lambda)(y - x)) + \lambda(1 - \lambda)C\|y - x\|^2$$

Therefore

$$\frac{f(y) - f(x)}{\|y - x\|} \leq \frac{f(x + (1 - \lambda)(y - x)) - f(x)}{(1 - \lambda)\|y - x\|} + \lambda C\|y - x\| \tag{2.12}$$

Let $\lambda \longrightarrow 1^-$ enough in (2.12) and using (2.11) we obtain

$$\frac{f(y) - f(x)}{\|y - x\|} \leq \frac{\langle \zeta, y - x \rangle}{\|y - x\|} + C\|y - x\|$$

It's become (2.10), and thus the proof is complete. □

As a corollary, it holds

Corollary 2.49: Let $f : \Omega \longrightarrow \mathbb{R}$ be semi-concave with semi-concavity C and $[x, y] \subset \Omega$. For $p \in D^+f(x)$ and $q \in D^+f(y)$, it holds

$$\langle q - p, y - x \rangle \leq 2C\|y - x\|^2$$

If f os semiconcave with semi-concavity C, then $g = -f$ is semi-convex with the same constant C. So we obtain

Proposition 2.50: Consider $\Omega \subset \mathbb{R}^n$ is an open set. Let $f : \Omega \longrightarrow \mathbb{R}$ is semi-convex with linear modulus C, then a vector $\zeta \in \mathbb{R}^n$ belong to $D^- f(x)$ if and only if

$$f(y) - f(x) - \langle \zeta, y - x \rangle \geq -C\|y - x\|^2 \tag{2.13}$$

for any point $y \in \Omega$ such that $[x, y] = \{\lambda x + (1 - \lambda y) : \lambda \in [0, 1]\} \subseteq \Omega$.

Proof. Since $-f$ is semi-concave with linear modulus C, from proposition 2.48 we obtain

$$\begin{aligned}
-\zeta \in D^+(-f)(x) &\Longleftrightarrow (-f)(y) - (-f)(x) - \langle -\zeta, y - x \rangle \leq C\|y - x\|^2 \\
\zeta \in D^- f(x) &\Longleftrightarrow -f(y) + f(x) + \langle \zeta, y - x \rangle \leq C\|y - x\|^2 \\
\zeta \in D^- f(x) &\Longleftrightarrow f(y) - f(x) - \langle \zeta, y - x \rangle \geq -C\|y - x\|^2
\end{aligned}$$

for any $y \in \Omega$ such that $[x, y] \subseteq \Omega$. □

From the definition of proximal subdifferentials and theorem 2.26, we see that in general $\partial_P f(x) \subset D^- f(x)$ but the converse is not true in general. The two notions coincide, however, for a semiconvex function with a linear modulus, by using proposition 2.50 above.

Proposition 2.51:

(i) Let $f : \Omega \longrightarrow \mathbb{R}$ is semi-convex with linear modulus C, then $\partial_P f(x) = D^- f(x)$.

(ii) Let $f : \Omega \longrightarrow \mathbb{R}$ is semi-cave with linear modulus C, then $\partial^P f(x) = D^+ f(x)$.

The converse is also true, i.e, if $f : \Omega \longrightarrow \mathbb{R}$ is continuous and $D^+ f(x) = \partial^P f(x)$, i.e there exists $C > 0$ such that

$$f(y) - f(x) \leq \langle \zeta, y - x \rangle + C\|y - x\|^2 \qquad \forall\, y \in \Omega \qquad \forall\, \zeta \in \partial^P f(x)$$

then clearly f is semi-concave with semi-concavity C. Indeed, for any vector h such that $[x - h, x + h] \subset \Omega$, for $\zeta \in \partial^P f(x)$, we have

$$\begin{aligned}
f(x + h) + f(x - h) - 2f(x) &= f(x + h) - f(x) + f(x - h) - f(x) \\
&= \langle \zeta, h \rangle + \langle \zeta, -h \rangle + 2C\|h\|^2 = 2C\|h\|^2
\end{aligned}$$

By proposition 2.35, we have f is semi-concave with semi-concavity C. So we summarize into

Proposition 2.52 (Properties of semi-concave functions)**:** Let $\Omega \subset \mathbb{R}^n$ be open and $f : \Omega \longrightarrow \mathbb{R}$ is continuous. Then the following are equivalent

(i) f is semi-concave with concavity C.

(ii) $f(y) - f(x) \leq \langle \zeta, y - x \rangle + C\|y - x\|^2$ for all $\zeta \in \partial^P f(x)$.

And the similar result for semi-convex function holds

Proposition 2.53 (Properties of semi-convex functions): Let $\Omega \subset \mathbb{R}^n$ be open and $f : \Omega \longrightarrow \mathbb{R}$ is continuous. Then the following are equivalent

(i) f is semi-convex with linear modulus C.

(ii) $f(y) - f(x) \geq \langle \zeta, y - x \rangle - C\|y - x\|^2$ for all $\zeta \in \partial_P f(x)$.

Finally, from proposition 2.31, we know that f is convex if and only if its epigraph is convex. We state the generalized version for semi-convex function. First from theorem 3.6.6 in [12] or [14], we have the following result.

Proposition 2.54: Assume $\Omega \subset \mathbb{R}^n$ is bounded, open and convex, $f : \Omega \longrightarrow \mathbb{R}^n$ is Lipschitz.

(i) f is semi-convex with linear modulus if and only if epi(f) has positive reach.

(ii) f is semi-concave with linear modulus if and only if hypo(f) has positive reach.

From this, we obtain

Theorem 2.55: Let $\Omega \subset \mathbb{R}^n$ is bounded, open and convex.

(i) A function $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ is semi-convex if and only if f is locally Lipschitz and epi(f) has positive reach.

(ii) A function $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ is semi-concave if and only if f is locally Lipschitz and hypo(f) has positive reach.

Proof. See [12] and [14]. It's just a corollary of proposition 2.54. If f is semi-convex then clearly f is locally Lipschitz by proposition 2.36. For each neighborhood Ω_x of x such that f is Lipschitz on Ω_x with some constant C_x, using proposition 2.54, we obtan epi(f) is proximally smooth, or it's has positive reach in a neighborhood of $(x, f(x))$. Since it's true for all $x \in \Omega$, we conclude epi(f) has positive reach. Here we already using the boundedness of Ω to obtain the minimum of reach(.).
For the converse, aslo choose a neighborhood Ω_x of x such that f is Lipschitz on this, and also using proposition 2.54 we obtain the semi-convexity of f on Ω_x, for all $x \in \Omega$, so f is semi-convex in Ω. Here we using the boundedness of Ω to obtain the maximal linear modulus of f. $\square$

We finish this section by the following theorem,

Theorem 2.56 (Alexandroff's theorem): Assume $\Omega \subset \mathbb{R}^n$ is open, and $f : \Omega \longrightarrow \mathbb{R}$ be semi-concave, then f is a.e twice differentiable in Ω, i.e for a.e $x \in \Omega$, there exists a vector $\zeta \in \mathbb{R}^n$ and a symmetric matrix B such that

$$\lim_{y \longrightarrow x} \frac{f(y) - f(x) - \langle \zeta, y - x \rangle + \langle B(y - x), y - x \rangle}{\|y - x\|^2} = 0$$

2.3 Linear control systems

2.3.1 Integral of vector–valued functions

Let $(E, \|.\|)$ is a Banach space and $u : [a,b] \longrightarrow E$ is continuous. A partition of $[a,b]$

$$P(a_0, a_1, \dots, a_n, c_1, \dots, c_n) \qquad \text{where} \qquad a = a_0 < c_1 < a_1 < \dots < c_n < a_n = b$$

Also define $|P| = \max\{a_{i+1} - a_i : i = \overline{0, n-1}\}$ and $\mathcal{P}([a,b])$ is the collection of all partitions of $[a,b]$. A Riemann sum with respect to the partition P is

$$S(u,P) = \sum_{i=1}^{n} u(c_i)(a_i - a_{i-1})$$

Theorem 2.57: Then there exists a unique $v \in E$ such that

$$\lim_{|P| \longrightarrow 0} S(u,P) = v$$

We call v is the integral of u over $[a,b]$, denoted by $\int_a^b u(t)\,dt$.

In case $E = \mathbb{R}^n$ and $u = (u_1, u_2, \dots, u_n)$, then the continuity of u is equivalent to the continuity of u_i for all $i = \overline{1,n}$. And it's not hard to see that the integral v in this case has the form

$$v = \int_a^b u(t)\,dt = \left(\int_a^b u_1(t)\,dt, \dots, \int_a^b u_n(t)\,dt \right)$$

Our goal is the following theorem, which is called the fundamental theorem of calculus for vector-valued function

Theorem 2.58: Assume that $f \in C([a,b], E)$, f is continuously differentiable on (a,b), i.e $f \in C^1((a,b), E)$ and f extends to a continuous function on $[a,b]$ which is still denoted by f. Then

$$f(b) - f(a) = \int_a^b f'(t)\,dt$$

We also have the usual property

$$\left\| \int_a^b u(t)\,dt \right\| \leq \int_a^b \|u(t)\|\,dt$$

2.3.2 The solution of linear control systems

Consider the linear ODE system

$$\begin{cases} x'(t) &= A \cdot x(t) \qquad \text{a.e} \quad t \in [0, +\infty) \\ x(0) &= x_0 \end{cases} \tag{2.14}$$

where $A \in M^{n\times n}(\mathbb{R})$, $x_0 \in \mathbb{R}^n$ is initial condition and $x(.) : [0,\infty]) \longrightarrow \mathbb{R}^n$ is a unknown function.
We call a function $x(.)$ which is absolutely continuous, $x(0) = 0$ and $x'(t) = A \cdot x(t)$ a.e in $[0,\infty)$ is a solution of (2.14).[1]

Now we have some observations. Assume $x(.)$ satisfies (2.14), then it must be continuous since it's has derivative a.e, and furthermore since

$$x'(t) = Ax(t)$$

we obtain that $x'(.)$ is also continuous, and similarly, we can see $x(.)$ is smooth, or $C^\infty(0,T)$. for any $T < \infty$. Thus by the fundamental theorem of Calculus, we obtain

$$x(t) - x(0) = \int_0^t x'(t)\,dt$$

Therefore (2.14) will has a solution of the form

$$x(t) = x_0 + \int_0^t A \cdot x(s)\,ds$$

We will prove that that there exist a unique solution like this. First, we introduce a classical representation of this solution, via the following terminology, since A can be though as a linear operator from $\mathbb{R}^n \longrightarrow \mathbb{R}^n$

Theorem 2.59: For $A \in L(H,H)$ where H is a Hilbert space, we have the following series is convergence in $L(H,H)$

$$e^A = \sum_{n=0}^{\infty} \frac{A^n}{n!} \qquad \text{where } A^n = A \circ A \circ \ldots \circ A \quad (n \text{ time})$$

and has some properties, if $A, B \in L(H,H)$ then

(a) If $A.B = B.A$ then $e^{A+B} = e^A.e^B = e^B.e^A$

(b) $e^0 = \text{Id}$ and $(e^A)^{-1} = e^{-A}$, $e^{(x+y)A} = e^{xA}.e^{yA}$.

(c) If $P \in M^{n\times n}(\mathbb{R})$ and $Q = P^{-1}AP$ then $e^Q = p^{-1}e^A P$.

(d) $\frac{d}{dt}e^{At} = Ae^{At}$.

It's easy to see if H is a Hilbert space then $L(H,H)$ is a Banach space with the usual sup-norm. So we can using the theory of integral of vector-valued function here to derive a nice property. For example, take $f : [0,T] \longrightarrow L(\mathbb{R}^n,\mathbb{R}^n)$ where

$$f(t) = e^{At} \in L(\mathbb{R}^n,\mathbb{R}^n) \qquad \text{with} \qquad f(t)(v) = e^{At} \cdot v$$

then from theorem 2.59 above we have f is continuously differentiable on $(0,T)$, f can be extended continuously on $[0,T]$ clearly, so by the Fundamental theorem of Calculus 2.58, we have

[1] The notion of absolutely continuous, see 2.3.7

Proposition 2.60: For any $t \in [0, T]$ we have

$$f(t) - f(0) = \int_0^t f'(s)\, ds \quad \Longrightarrow e^{tA} - 1 = \int_0^t A e^{sA}\, ds$$

Recall a useful fact

Theorem 2.61: [Gronwall's inequality] Let φ continuous $[a, b]$, $\varphi \geq 0$, assume there exists constants K, C s.t

$$0 \leq \varphi(t) \leq K + C \int_a^t \varphi(s)\, ds \qquad \forall\, t \in [a, b]$$

then

$$0 \leq \varphi(t) \leq K e^{c(t-a)} \qquad \forall\, t \in [a, b]$$

Proof. Let $F(t) = \int_0^t \varphi(s)\, ds$ for $t \in [a, b]$, we have $F'(t) = \varphi(t) \geq 0$ foe all $t \in [a, b]$, so F is increase so $F(t) \geq F(a) = 0$. So for all $t \in [a, b]$

$$\varphi(t) \leq K + CF(t) \Longrightarrow \varphi(t) - CF(t) \leq K \Longrightarrow \left(F(t)e^{-Ct}\right)' \leq Ke^{-Ct} \Longrightarrow F(s)e^{-Cs} \leq K \int_a^s e^{-Ct}\, dt$$

Hence

$$F(s) \leq \frac{K}{C}\left(e^{C(s-a)} - 1\right) \Longrightarrow \varphi(s) \leq K + CF(s) \leq Ke^{C(s-a)}$$

□

From this fact, we easily conclude that $x(t) = e^{At}x_0$ is the unique solution of (2.14).

Theorem 2.62: The linear system (2.14) has the unique solution of the form $x(t) = e^{At}x_0$

Proof. Clearly $x(t) = e^{At}x_0$ satisfies the system (2.14), furthermore, it's is smooth, thus it's a solution. Now we will prove that it's unique by using the Gronwall's inequality 2.61.

Assume $u(.), v(.)$ are two solution of (2.14), then

$$\begin{cases} (u-v)'(t) &= A \cdot (u-v)(t) \\ (u-v)(t) &= 0 \end{cases}$$

And thus, the function $\varphi = u - v$ is satisfies

$$\begin{cases} \varphi'(t) &= A \cdot \varphi(t) \qquad \forall\, t > 0 \\ \varphi(0) &= 0 \end{cases}$$

Clearly $\varphi'(.)$ is continuous, so by the fundamental theorem of Calculus we obtain

$$\varphi(t) = \int_0^t \varphi'(s)\, ds = \int_0^t A \cdot \varphi(s)\, ds \Longrightarrow \|\varphi(t)\| \leq \left\| \int_0^t A \cdot \varphi(s)\, ds \right\| \leq \int_0^t \|\varphi(s)\|\, ds$$

Now using Gronwall's inequality 2.61, we obtain $\varphi(.) \equiv 0$, thus $u \equiv v$ and the solution is unique. □

Now applying to our problem, we consider the linear control system

$$\begin{cases} x'(t) &= Ax(t) + Bu(t) \qquad t > 0 \\ x(0) &= x_0 \end{cases} \tag{2.15}$$

where $A \in \mathbb{M}^{n\times n}$, $B \in \mathbb{M}^{n\times m}$, $1 \le m \le n$ and $U = [-1,1]^m \subset \mathbb{R}^m$. The set of admissible control is

$$\mathcal{U}_{ad} = \{u : [0,\infty) \longrightarrow U : u \text{ is measurable}\}$$

The most difficulty in this case it that, the function $t \longmapsto Ax(t) + Bu(t)$ is not continuous in general, so we can not using the fundamental theorem of Calculus as usual in here. But by imitating the above process, we can claim that, if (2.15) has solution, it's must be unique.

Proposition 2.63: If $u(.), v(.)$ are two solution of (2.15) in time $[0,T]$ for $T < \infty$. Then $u \equiv v$ in $[0,T]$.

Proof. We also obtain

$$\begin{cases} (u-v)'(t) &= A\cdot(u-v)(t) \\ (u-v)(t) &= 0 \end{cases}$$

and thus $u \equiv v$ by using Gronwall's inequality because $(u-v)'(.)$ is continuous. □

We now present an other approach. Assume $x(.)$ is a solution on $[0,T]$ of (2.15), then for $t \in (0,T)$ we have

$$\begin{aligned} x'(t) = A\cdot x(t) + Bu(t) &\Longrightarrow -Ae^{-At}x(t) + e^{-At}x(t) = e^{-At}Bu(t) \\ &\Longrightarrow \left(e^{-At}x(t)\right)'(t) = e^{-At}Bu(t) \end{aligned}$$

Clearly $s \longmapsto e^{-sA}Bu(s)$ is $L^1(0,t)$ in this case since $\|u(t)\|$ is bounded. So we guest that

$$e^{-At}x(t) - x_0 = \int_0^t e^{-As}Bu(s)\,ds \qquad \text{i.e,} \qquad x(t) = e^{tA}x_0 + \int_0^t e^{(t-s)A}Bu(s)\,ds$$

Now define

$$y^{x_0,u}(t) = e^{At}x_0 + \int_0^t e^{A(t-s)}Bu(s)\,ds$$

We can see $x(.) = y^{x_0,u}(.)$ is continuous, and $x(0) = x_0$. The rest is check that $x(.)$ satisfies $x'(t) = Ax(t) + Bu(t)$. Setting

$$g(t) = \int_0^t e^{-sA}Bu(s)\,ds \tag{2.16}$$

It's easy to check that g is continuous. Assume g is differentable and $g'(t) = e^{-tA}Bu(t)$, then

$$\begin{aligned} x'(t) &= Ae^{At}x_0 + Ae^{At}g(t) + e^{At}g'(t) \\ &= Ax(t) + e^{At}\cdot e^{-At}Bu(t) = Ax(t) + Bu(t) \end{aligned}$$

But again, $s \longmapsto e^{-As}Bu(s)$ is not continuous, so it's not straight-forward to obtain this fact about g. Thus, our work is now prove g is differentiable and $g'(t) = e^{-tA}Bu(t)$. Define $f(s) = e^{-sA}Bu(s)$, then $f \in L^1((0,t),\mathbb{R}^n)$. Let's compute

$$\frac{g(t+h)-g(t)}{h} = \frac{1}{h}\int_t^{t+h} f(s)\,ds$$

Now writing $f = (f_1, f_2, \dots, f_n)$, we obtain $f_i \in L^1((0,t),\mathbb{R})$ for all $i \in \overline{1,n}$, and

$$\frac{g(t+h)-g(t)}{h} = \frac{1}{h}\int_t^{t+h} f(s)\,ds = \left(\frac{1}{h}\int_t^{t+h} f_1(s)\,ds, \dots, \frac{1}{h}\int_t^{t+h} f_n(s)\,ds\right)$$

Thus, our goal is prove that for any $i \in \overline{1,n}$ then

$$\lim_{h\longrightarrow 0}\left(\frac{1}{h}\int_t^{t+h} f_i(s)\,ds\right) = f(t) \Longleftrightarrow \lim_{h\longrightarrow 0}\left(\frac{1}{h}\int_t^{t+h}\Big(f_i(s) - f(t)\Big)\,ds\right) = 0 \tag{2.17}$$

for any $i \in \overline{1,n}$. Recall the Lebesgue's differentiation theorem

Theorem 2.64: Assume $\Omega \subset \mathbb{R}^n$ is open and $f \in L^1_{\text{loc}}(\Omega,\mathbb{R})$, then

$$\lim_{r\longrightarrow 0}\frac{1}{\mu(B(x,r))}\int_{B(x,r)} |f(y) - f(x)|\,dy = 0 \qquad \text{for a.e } x \in \Omega$$

where μ denote the Lebesgue measure on $\mathbb{R}^n$.

For $f_i \in L^1((0,T),\mathbb{R})$, $i \in \overline{1,n}$, note that in $\mathbb{R}$, $\mu(B(t,r)) = \mu(t-r,t+r) = 2r$, we have

$$\left|\frac{1}{h}\int_t^{t+h}\Big(f_i(s) - f(t)\Big)\,ds\right| \le \frac{1}{h}\int_{(t,t+h)}\big|f_i(s) - f_i(t)\big|\,ds = \frac{2}{\mu(B(t,h))}\int_{B(t,h)}\big|f_i(s) - f_i(t)\big|\,ds$$

Using Lebesgue's differentiation theorem, we obtain for a.e $t \in (0,T)$

$$\lim_{h\longrightarrow 0}\left(\frac{2}{\mu(B(t,h))}\int_{B(t,h)}\big|f_i(s) - f_i(t)\big|\,ds\right) = 0$$

and thus for a.e $t \in (0,T)$.

$$\lim_{h\longrightarrow 0}\left(\frac{1}{h}\int_t^{t+h}\Big(f_i(s) - f(t)\Big)\,ds\right) = 0$$

Thus, (2.17) is true for a.e $t \in (0,T)$.

In summarize, we conclude that given $u \in \mathcal{U}_{ad}$, the linear control system (2.15) has a unique solution given by

$$x(t) = y^{x_0,u}(t) = e^{At}x_0 + \int_0^t e^{A(t-s)}Bu(s)\,ds \tag{2.18}$$

Recall that for any continuous function $f : [a,b] \longrightarrow \mathbb{R}^n$, define

$$\begin{aligned} F : [a,b] &\longrightarrow \mathbb{R}^n \\ t &\longmapsto \int_a^t f(s)\, ds \end{aligned}$$

then F is (Fréchet) differentiable in (a,b), continuous on $[a,b]$ and furthermore

$$F'(t) = f(t) \qquad \forall\, t \in (a,b) \tag{2.19}$$

Also the converse is true, if $F : [a,b] \longrightarrow \mathbb{R}^n$ is differentiable on (a,b) and $F' = f$ is continuous on $[a,b]$, then (2.19) holds.

The linear control system (2.15) give us that for a control $u \in \mathcal{U}_{ad}$ and $T < \infty$, the trajectory $\mathbf{x}(.) = y^{x_0,u}(.)$ is continuous on $[0,T]$, furthermore it's differentiable on $(0,T)$ where

$$\mathbf{x}'(t) = A\mathbf{x}(t) + Bu(t)$$

But $\mathbf{x}'(t)$ is not continuous in general, so it isn't straight-forward to obtain the formula

$$\mathbf{x}(t) - \mathbf{x}(0) = \int_0^t \mathbf{x}'(s)\, ds \tag{2.20}$$

But in this case, by using the uniqueness of solution of (2.15), we can obtain this fact. Indeed, since $u : [0,T] \longrightarrow [-1,1]^m$ is bounded, and $x : [0,T] \longrightarrow \mathbb{R}^n$ is continuous, and $[0,T]$ is compact, we must have $\|x(.)\|$ is bounded on $[0,T]$, thus

$$x'(.) = A \cdot x(.) + Bu(.) \Longrightarrow x'(.) \in L^1((0,T), \mathbb{R}^n)$$

Setting

$$h : [0,T] \longrightarrow \mathbb{R}^n \qquad \text{where} \qquad t \longmapsto x_0 + \int_0^t x'(s)\, ds$$

Writing $x = (x_1, x_2, \ldots, x_n)$, then $x' = (x_1', \ldots, x_n')$. So if we assume $h = (h_1, h_2, \ldots, h_n)$ then from definition of h we obtain

$$h_i(t) = \mathrm{pr}_i(x_0) + \int_0^t x_i'(s)\, ds$$

where $\mathrm{pr}_i : (a_1, \ldots, a_n) \longmapsto a_i$ is the i^{th} projection. Since every x_i is belong to $L^1((0,T), \mathbb{R})$, using Lebesgue differentiation theorem 2.64 we obtain

$$\lim_{\varepsilon \longrightarrow 0} \frac{h_i(t+\varepsilon) - h_i(t)}{\varepsilon} = \lim_{\varepsilon \longrightarrow 0} \left(\frac{1}{\varepsilon} \int_t^{t+\varepsilon} x_i'(s)\, ds \right) = x_i'(t) \qquad \text{for a.e } t \in (0,T)$$

Therefore we have h is differentiable and $h'(t) = x'(t)$ for a.e $t \in (0,T)$. Also $h(0) = x_0$ by the definition of h, so h is another solution of the linear system (2.15). Now by the uniqueness of the solution, we obtain $h \equiv y^{x_0,u}$, thus

$$x(t) = y^{x_0,u}(t) = h(t) = x_0 + \int_0^t x'(s)\, ds \qquad \text{for a.e } t \in (0,T)$$

Later we will observe that $x(.) = y^{x_0,u}$ is Lipschitz on every compact interval $[0,T]$, and hence it's absolutely continuous on $[0,T]$.

2.3.3 Example: Rocket rail road car

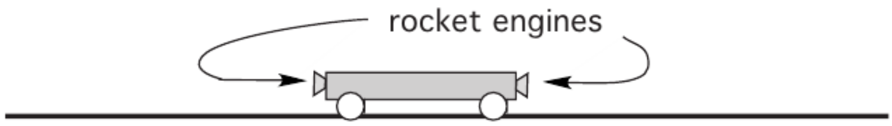
We introduce an example before going to state main problem

Example. [The rocket car] Imaging a rail road car powered by rocket engines on each side. We introduce the variables

- $x(t)$ is the position of the rocket rail road car on the train track at time t.
- $v(t)$ is the velocity of the rocket rail road car at time t.
- $F(t)$ is the force from the rocket engines at time t.

where we only consider $F(t) \in [-1, 1]$ and the sign of $F(t)$ depends on which engine is firing.

Our goal. Construct $F(.)$ in order to drive the rocket rail road car to the origin 0 with zero velocity in a minimum amount of time.



A ROCKET CAR ON A TRAIN TRACK

Mathematical Model. Assuming that the rocket railroad car has mass $m = 1$, the motion law is

$$mx''(t) = F(t) \Longleftrightarrow x''(t) = \frac{F(t)}{m} = u(t)$$

where $u(.)$ is understood as a control function. So the motion equation of the rocket railroad car is

$$\begin{cases} x''(t) & = u(t) \\ x(0) & = x_0 \qquad \text{and} \qquad v(0) = v_0 \end{cases} \tag{2.21}$$

where $u(t) \in U = [-1, 1]$ for all $t \geq 0$, x_0 is the position of the rocket car at time 0 and v_0 is the velocity of the rocker railroad car at x_0. By setting

$$z(t) = \begin{pmatrix} x(t) \\ v(t) \end{pmatrix} \qquad A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

we can rewrite (2.21) as first order control system

$$\begin{cases} z'(t) & = A \cdot z(t) + b \cdot u(t) \\ z(0) & = (x_0, v_0)^T \end{cases}$$

A geometric solution We will introduce a way to steer (x_0, v_0) to the origin (0,0). First of all, let us guess that to find an optimal control solution we will need only to consider cases $u = 1$ or $u = -1$. In other words, we will focus our attention only upon those controls for which at each moment of time either the left or the right rocket engine is fired at full power. We will see later that this assumption is correct.

Case $u = 1$**.** Assume $u = 1$ on some time interval, during which, from (2.21) we have

$$\begin{cases} x'(t) & = v(t) \\ v'(t) & = 1 \end{cases} \Longrightarrow v(t)v'(t) = x'(t) \Longrightarrow \frac{1}{2}(v^2(t))' = x'(t)$$

Now let t_0 belong to this the interval where $u = 1$ and integrate from t_0 to t, where t so small such that $[t_0, t]$ also belong to this time interval

$$\frac{v^2(t)}{2} - \frac{v^2(t_0)}{2} = x(t) - x(t_0) \Longrightarrow v^2(t) = 2x(t) + \underbrace{\left(v^2(t_0) - 2x(t_0)\right)}_{b_0} \tag{2.22}$$

In other words, in the time interval such that $u = 1$, the trajectory stays on the curve $v^2(t) = 2x(t) + b_0$, where $b_0 = v^2(t_0) - 2x(t_0)$ is a constant respect to the initial condition (x_0, v_0).

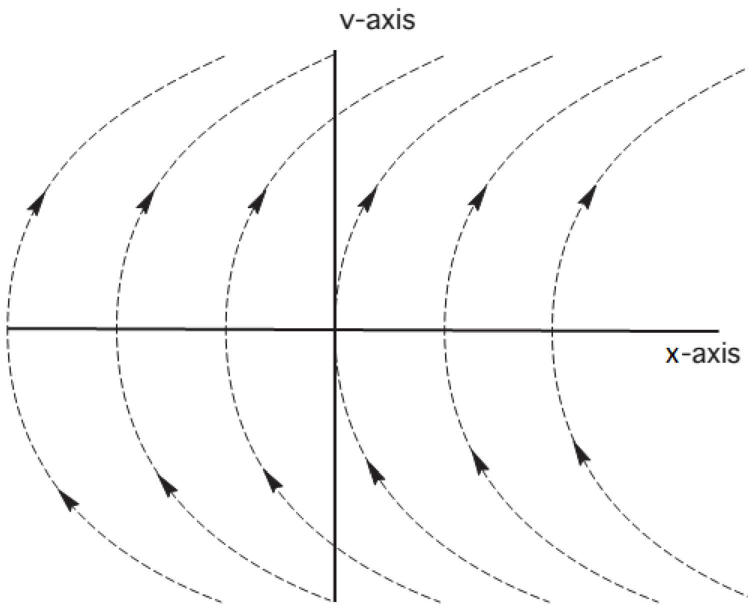


Figure 1: The curves $v^2(t) = 2x(t) + b_0$ in case $u = 1$.

Case $u = -1$**.** Assume $u = -1$ on some time interval, during which, from (2.21) we have

$$\begin{cases} x'(t) & = v(t) \\ v'(t) & = -1 \end{cases} \Longrightarrow v(t)v'(t) = -x'(t) \Longrightarrow \frac{1}{2}(v^2(t))' = -x'(t)$$

Now let t_0 belong to this the interval where $u = -1$ and integrate from t_0 to t, where t so small such that $[t_0, t]$ also belong to this time interval

$$\frac{v^2(t)}{2} - \frac{v^2(t_0)}{2} = -x(t) + x(t_0) \Longrightarrow v^2(t) = 2x(t) + \underbrace{\left(v^2(t_0) + 2x(t_0)\right)}_{c_0} \tag{2.23}$$

In other words, in the time interval such that $u = 1$, the trajectory stays on the curve $v^2(t) = 2x(t) + b_0$, where $c_0 = v^2(t_0) + 2x(t_0)$ is a constant respect to the initial condition (x_0, v_0).

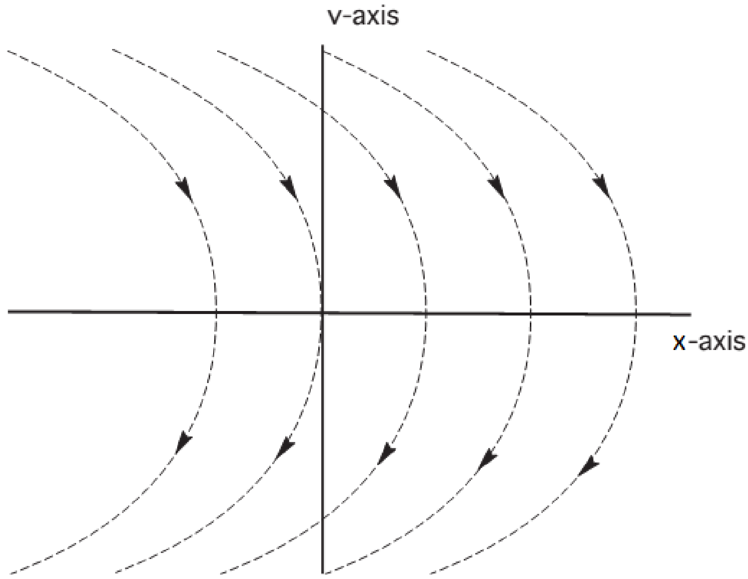


Figure 2: The curves $v^2(t) = -2x(t) + c_0$ in case $u = -1$.

Geometric Interpretation Now, conclude from two cases above and equations (2.22), (2.23), we can design an control u', which steers the initial data (x_0, v_0) (we denote by the black dot) to the origin. For example, assume it live on the plane like following picture, it's easy to see that if we set $u' = -1$, causing this point move down along the parabola of $u = -1$ and pass through $(x, 0, v_0)$, then when it first meets the parabola of $u = 1$ and pass through the origin, then switch $u' = 1$ on this. This control u' is steer (x_0, v_0) to $(0, 0)$ by switching only one time.

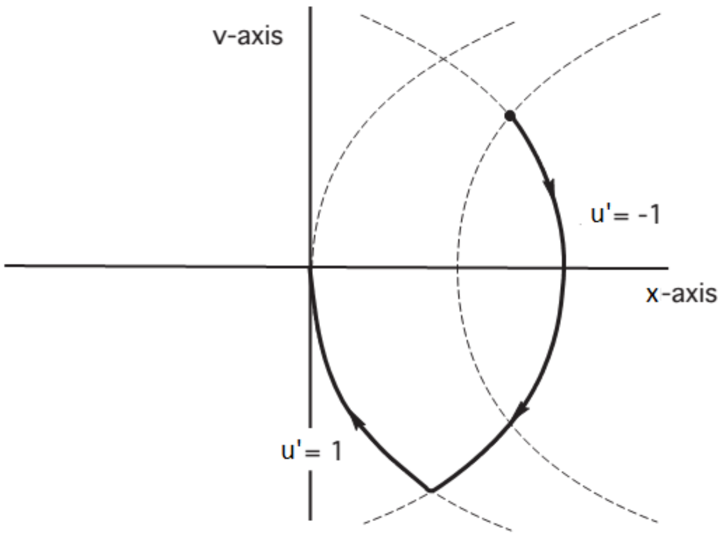


Figure 3: The trajectory causing by the control u' move (x_0, v_0) to the origin

We shall see later that such a control like u' is optimal.

2.3.4 Example: Harmonic oscillation

Now we consider the simple harmonic oscillation of an oscillating weight (of unit mass, i.e $m = 1$) in a spring, with the control is interpreted as an exterior force acting on this oscillating weight.

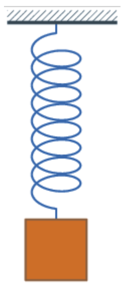


Figure 4: Vibrating Spring

For simplify, assume some conditions lead to consider its equation

$$x'' + x = u$$

where $x(t)$ denote the position oat time t, $v(t) = x'(t)$ denote the velocity at time t, $a(t) = x''(t)$ denote the acceleration of this oscillating weight at time t. And u be the exterior force acting to this mass hanging from a spring.

Our goal. Given the initial position and the initial velocity, we want to construct an optimal exterior forcing α that brings the motion to a stop in minimum time.

Mathematical Model. Considering the equation $x'' + x = u$, where $u : \mathbb{R} \longrightarrow U = [-1, 1]$ and $x_0 = x(0), v_0 = v(0)$ be the given initial data. The motion equation is $v'(t) + x(t) = u(t)$, i.e

$$\begin{cases} x'(t) &= v(t) & x(0) = x_0 \\ v'(t) &= -x(t) + u(t) & v(0) = v_0 \end{cases} \tag{2.24}$$

By setting

$$z(t) = \begin{pmatrix} x(t) \\ v(t) \end{pmatrix} \qquad A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

We have that

$$z'(t) = \begin{pmatrix} v(t) \\ v'(t) \end{pmatrix} \qquad z(0) = \begin{pmatrix} x_0 \\ v_0 \end{pmatrix} \Longrightarrow \begin{cases} z'(t) &= A \cdot z(t) + b \cdot u(t) \\ z(0) &= (x_0, v_0)^T = z_0 \end{cases}$$

This become a first order control system as we have known in general.

2.3.5 Basic notions and properties

Consider the linear control system

$$\begin{cases} x'(t) &= Ax(t) + Bu(t) \qquad t > 0 \\ x(0) &= x_0 \end{cases} \tag{2.25}$$

where $A \in \mathbb{M}^{n\times n}$, $B \in \mathbb{M}^{n\times m}$, $1 \leq m \leq n$ and $U = [-1, 1]^m \subset \mathbb{R}^m$. The set of admissible control is

$$\mathcal{U}_{ad} = \{u : [0, \infty) \longrightarrow U : u \text{ is measurable}\} \tag{2.26}$$

It's easy to see some basic properties of $\mathcal{U}_{ad}$

Theorem 2.65: The set $\mathcal{U}_{ad}$ is convex, symmetric in $L^\infty([0, +\infty), \mathbb{R}^m)$

Proof. For the convex, if $u, v \in \mathcal{U}_{ad}$ then for all $\lambda \in (0, 1)$, clearly $w = \lambda u + (1-\lambda)v \in \mathcal{U}_{ad}$ since it's measurable from $[0, +\infty) \longrightarrow U$. It's also obvious that if $u \in \mathcal{U}_{ad}$ then $-u \in \mathcal{U}_{ad}$ so $\mathcal{U}_{ad}$ is symmetric. Note that $U = [-1, 1]^m$ in this case is convex. $\square$

Given $u \in \mathcal{U}_{ad}$, the trajectory starting from x_0 with the control u can be presented as

$$y^{x_0,u}(t) = e^{tA}x_0 + \int_0^t e^{(t-s)A}Bu(s)\, ds \tag{2.27}$$

Let's define the **reachable set for time** t to be

$$\mathcal{R}(t) = \{x_0 \in \mathbb{R}^n : \exists\, u \in \mathcal{U}_{ad} \text{ such that} : x(t) = 0\}$$

And define the **overall reachable set** to be

$$\mathcal{R} = \bigcup_{t\geq 0} \mathcal{R}(t)$$

Using this expression, we can assert some basic properties of the reachable set $\mathcal{R}(t)$.

Theorem 2.66: For a fixed time $t > 0$, we have

(i) $\mathcal{R}(t)$ is convex, symmetric and compact.

(ii) $\mathcal{R}(t) \subset \mathcal{R}(t')$ for all $t' \geq t \geq 0$.

(iii) $\mathcal{R}$ is convex and symmetric.

Proof.

(i) For the convex, assume that $x_0, y_0 \in \mathcal{R}(t)$, i.e there exist controls $u, v \in \mathcal{U}_{ad}$ s.t

$$x_0 = -\int_0^t e^{-sA}Bu(s)\,ds \qquad y_0 = -\int_0^t e^{-sA}Bv(s)\,ds \Longrightarrow z = -\int_0^t e^{-sA}Bw(s)\,ds$$

where $\lambda \in (0,1)$ and $z = \lambda x_0 + (1-\lambda)y_0$ and $w(.) = \lambda u(.) + (1-\lambda)v(.) \in \mathcal{U}_{ad}$ by the convexity of $\mathcal{U}_{ad}$. So $w \in \mathcal{R}(t)$, i.e $\mathcal{R}(t)$ is convex. For the symmetric, similarly

$$x_0 = -\int_0^t e^{-sA}Bu(s)\,ds \Longrightarrow -x_0 = -\int_0^t e^{-sA}Bw(s)\,ds$$

where $w(.) = -u(.) \in \mathcal{U}_{ad}$ since $\mathcal{U}_{ad}$ is symmetric, so $\mathcal{R}(t)$ is symmetric. For the compactness, let $(x_n) \subset \mathcal{R}(t)$ which are respect to the sequence of controls $(u_n) \subset U_{ad}$, where

$$x_n = -\int_0^t e^{-sA}Bu_n(s)\,ds \qquad \forall\, n \in \mathbb{N}$$

Then since $\{u_n\}$ is a bounded sequence in $L^\infty((0,t),\mathbb{R}^n)$, by theorem (2.5), there exists a subsequence $(u_{n_k}) \subset (u_n)$ s.t $u_{n_k} \xrightharpoonup{*} \overline{u} \in L^\infty((0,t),\mathbb{R}^n)$ in the weak* topology $\sigma(L^\infty, L^1)$, and easily seen that $\overline{u} \in \mathcal{U}_{ad}$. Hence we have

$$x_{n_k} = -\int_0^t (e^{-sA}B)\cdot u_{n_k}(s)\,ds \longrightarrow -\int_0^t (e^{-sA}B)\cdot \overline{u}(s)\,ds = \overline{x} \quad \text{as} \quad k \longrightarrow \infty$$

since $v(s) = e^{-sA}B : [0,t] \longrightarrow \mathbb{R}^m$ is integrable by

$$\int_0^t |v(s)|\,ds \leq \int_0^t e^{t\|A\|}\|B\|\,ds \leq te^{t\|A\|}\|B\| < \infty$$

So $x_{n_k} \longrightarrow \overline{x}$ in $\mathbb{R}^n$ and clearly $\overline{x} \in \mathcal{R}(t)$. So $\mathcal{R}(t)$ is compact in $\mathbb{R}^n$.

(ii) If $x \in \mathcal{R}(t)$ which respect to the control u, define $v = u\chi_{[0,t]} + 0.\chi_{[t,t']}$ where $t > t'$, then clearly $v \in \mathcal{U}_{ad}$ and $x \in \mathcal{R}(t')$ which respect to control v, hence $\mathcal{R}(t) \subset \mathcal{R}(t')$.

(iii) It's clearly from (i) and (ii).

$\square$

As an useful remark, we have $x_0 \in \mathcal{R}(t)$ iff exist $u(.) \in \mathcal{U}_{ad}$ such that

$$0 = e^{tA}x_0 + \int_0^t e^{(t-s)A}Bu(s)\,ds \Longleftrightarrow x_0 = -\int_0^t e^{-sA}Bu(s)\,ds$$

2.3.6 The minimum time function

We will study the continuity of the minimum time function for the linear control system (2.25), recall that given $u \in \mathcal{U}_{ad}$, the trajectory starting from x_0 with the control u can be presented as

$$y^{x_0,u}(t) = e^{tA}x_0 + \int_0^t e^{(t-s)A}Bu(s)\,ds$$

Now let's define the minimum time function, for fixed $x_0 \in \mathbb{R}^n \backslash \{0\}$, let $T(x)$ be the minimum amount of time to reach to the target S from x_0, i.e

$$T(x_0) = \inf_{u \in \mathcal{U}_{ad}} \{t > 0 : y^{x_0,u}(t) = 0\} \tag{2.28}$$

From this definition, one can see that

- $T(x_0)$ is finite for all $x_0 \in \mathcal{R}$
- $T(x_0) = +\infty$ for all $x_0 \in \mathbb{R}^n \backslash \mathcal{R}$.

Theorem 2.67 (Existence of the time–optimal control): For every $x_0 \in \mathcal{R}$, we have that

$$T(x_0) = \min_{u \in \mathcal{U}_{ad}} \{t > 0 : y^{x_0,u}(t) = 0\} \tag{2.29}$$

It means that there exists an optimal control $u^* \in \mathcal{U}_{ad}$ such that $y^{x_0,u^*}(T(x_0)) = 0$.

Proof. Since $x_0 \in \mathcal{R}$, we have that $T = T(x_0) < \infty$, so there exists a sequence of admissible control $\{u_k\} \subset \mathcal{U}_{ad}$ and an decreasing sequence $\{t_k\}$ converging to $T(x_0)$ such that

$$y^{x_0,u_k}(t_k) = 0 \qquad \forall\, k \in \mathbb{N}$$

Note that for each time t_k, we only consider the control $u_{t_k}(s)$ for $s \in [0, t_k]$, so we can assume that $u_{t_k}(s) \equiv 0$ for $s \geq t_k$.
Since $\|u_k\|_{L^\infty(0,t_1)} \leq 1$, there exists a subsequence $\{u_{k_l}\}$ such that $u_{k_l} \overset{*}{\rightharpoonup} \bar{u}$ weakly in the weak* topology $\sigma(L^\infty([0,t_1]), L^1([0,t_1])$, and clearly $\bar{u} \in \mathcal{U}_{ad}$. We claim that $y^{x_0,\bar{u}}(T) = 0$. Indeed we will prove that $y^{x_0,u_{k_l}}(t_{k_l}) \longrightarrow y^{x_0,\bar{u}}(T)$ as $l \longrightarrow \infty$, note that

$$y^{x_0,u_{k_l}}(t_{k_l}) = y^{x_0,u_{k_l}}(t_1)$$

since $u_{k_l}(s) = 0$ for all $s \geq t_{k_l}$, and $t_1 \geq t_{k_l}$ for all $l \in \mathbb{N}$. Now by weakly convergence, clearly

$$y^{x_0,u_{k_l}}(t_1) \longrightarrow y^{x_0,\bar{u}}(t_1) \qquad \text{as} \qquad l \longrightarrow \infty \tag{2.30}$$

We claim that $y^{x_0,\overline{u}}(t_1) = y^{x_0,\overline{u}}(T)$ by proving $\overline{u} = 0$ a.e in $[T, t_1]$. Indeed, since $u_{k_l} \overset{*}{\rightharpoonup} \overline{u}$ in $\sigma(L^\infty, L^1)$, so for any integrable function $g(.) : [0, t_1] \longrightarrow \mathbb{R}^m$ we have that

$$\int_0^{t_1} g(s) \cdot u_{k_l}(s)\, ds \longrightarrow \int_0^{t_1} g(s) \cdot \overline{u}(s)\, ds \qquad \text{as} \qquad l \longrightarrow \infty$$

Now take

$$g(s) = \begin{cases} \frac{\overline{u}(s)}{\|\overline{u}(s)\|} \cdot \chi_{[T,t_1]} & \text{if} \quad \overline{u}(s) \neq 0 \\ 0 & \text{if} \quad \overline{u}(s) = 0 \end{cases}$$

which is clearly integrable, and so we have that $u_{k_l}(s) = 0$ for $s \geq t_{k_l}$, so

$$\int_0^{t_1} g(s) \cdot u_{k_l}(s)\, ds = \int_T^{t_1} g(s) \cdot u_{k_l}(s)\, ds \longrightarrow \int_0^{t_1} g(s) \cdot \overline{u}(s)\, ds = \int_T^{t_1} \|\overline{u}(s)\|\, ds \qquad (2.31)$$

On the other hand, we have that

$$\left| \int_T^{t_1} g(s) \cdot u_{k_l}(s)\, ds \right| = \left| \int_T^{t_{k_l}} g(s) \cdot u_{k_l}(s) \right| ds \leq \int_T^{t_{k_l}} \|g(s)\| \cdot \|u_{k_l}(s)\|\, ds \longrightarrow 0 \qquad (2.32)$$

as $l \longrightarrow \infty$ since $t_{k_l} \longrightarrow T$. Compile (2.31) and (2.32) we conclude that

$$\int_T^{t_1} \|\overline{u}(s)\|\, ds = 0 \Longrightarrow \overline{u}(s) = 0 \quad \text{a.e in } [T.t_1] \qquad (2.33)$$

So from (2.30) and (2.33) we have that $y^{x_0,\overline{u}}(t_1) = y^{x_0,\overline{u}}(T)$, i.e

$$y^{x_0,\overline{u}}(T) = \lim_{l \longrightarrow \infty} y^{x_0,u_{k_l}}(t_1) = \lim_{l \longrightarrow \infty} y^{x_0,u_{k_l}}(t_{k_l}) = 0$$

The proof is complete. □

Theorem 2.68 (Dynamics programming principle for the minimum time function in linear case)**:** Let $x \in \mathcal{R}$, then it holds

$$\min_{u \in \mathcal{U}_{ad}} \{t + T\,(y^{x,u}(t))\} = T(x) \qquad \forall\; 0 < t < T(x)$$

Proof. Fix $x_0 \in \mathcal{R}$ and define $T = T(x_0) < \infty$, we also fix time $t \in (0, T)$. Let's prove by some steps.

1. $T(x_0) \leq \inf\{t + T\,(y^{x_0,u}(t)) : u \in \mathcal{U}_{ad}\}$
   Pick $u \in \mathcal{U}_{ad}$ arbitrary, we need to show that $T(x_0) \leq t + T(y^{x_0,u}(t))$. Indeed, let $x_u(.) = y^{x_0,u}(.)$ be the trajectory which is respect to u, i.e

$$\begin{cases} x_u'(s) & = Ax_u(s) + Bu(s) \qquad s > 0 \\ x(0) & = x_0 \end{cases}$$

   Now define $x_1 = x_u(t) = y^{x_0,u}(t)$, from (2.29), there exists an admissible control $\overline{u}$ such that

$$y^{x_1,\overline{u}}(T(x_1)) = 0 \qquad (2.34)$$

Define $x_{\overline{u}}(.) = y^{x_1,\overline{u}}(.)$ be the trajectory which respect to $\overline{u}$, i.e

$$\begin{cases} x'_{\overline{u}}(s) &= Ax_{\overline{u}}(s) + B\overline{u}(s) \qquad s > 0 \\ x_{\overline{u}}(0) &= x_1 \end{cases}$$

Define a new control u^* by this way

$$\begin{cases} u^*(s) &= u(s) \qquad \forall\, s \in (0,t) \\ u^*(s) &= \overline{u}(s) \qquad \forall\, s \in (t,T) \end{cases}$$

Then, let $x_{u^*}(.) = y^{x_0,u^*}(.)$ then

$$\begin{cases} x'_{u^*}(s) &= Ax_{u^*}(s) + Bu^*(s) \\ x_{u^*}(0) &= x_0 \\ x_{u^*}(t) &= x_1 \end{cases} \Longrightarrow \begin{cases} x_{u^*}(s) &= x_u(s) \qquad \forall\, s \in (0,t) \\ x_{u^*}(s) &= x_{\overline{u}}(s) \qquad \forall\, s \in [t,T) \end{cases}$$

Under the affect of u^*, we have $x_0 \longmapsto x_1 \longmapsto 0$, and since (2.34) we also have

$$y^{x_0,u^*}(t + T(x_1)) = 0 \Longrightarrow T(x_0) \le t + T(x_1) = t + T\left(y^{x_0,u}(t)\right)$$

so we obtain $T(x_0) \le \inf\{t + T\left(y^{x_0,u}(t)\right) : u \in \mathcal{U}_{ad}\}$ since u is choosing arbitrary.

2. $T(x_0) \ge \inf\{t + T\left(y^{x_0,u}(t)\right) : u \in \mathcal{U}_{ad}\}$
We only need to find some control $v \in \mathcal{U}_{ad}$ such that $T(x_0) \ge t + T(y^{x_0,v}(t))$. Indeed, from (2.29) there exists a control $u \in \mathcal{U}_{ad}$ such that $y^{x_0,u}(T(x_0)) = 0$. Let $x_u(.) = y^{x_0,u}(.)$ we have

$$\begin{cases} x'_u(s) &= Ax_u(s) + Bu(s) \qquad s > 0 \\ x(0) &= x_0 \end{cases}$$

Now define $x_1 = x_u(t) = y^{x_0,u}(t)$, and a new control $v \in \mathcal{U}_{ad}$ by

$$v(s) := \begin{cases} u(s+t) & 0 \le s \le T - t \\ 0 & t < s \le T(x) \end{cases}$$

and one can see that $y^{x_1,v}(T - t) = 0$, indeed let's recall that

$$y^{x_0,u}(t) = e^{tA}x_0 + \int_0^t e^{(t-s)A}Bu(s)\, ds = x_1$$

So we have

$$\begin{aligned} y^{x_1,v}(T-t) &= e^{(T-t)A}x_1 + \int_0^{T-t} e^{(T-t-s)A}Bv(s)\, ds \\ &= e^{(T-t)A}\left(e^{tA}x_0 + \int_0^t e^{(t-s)A}Bu(s)\, ds\right) + \int_0^{T-t} e^{(T-t-s)A}Bv(s)\, ds \\ &= e^{(T-t)A}\left(e^{tA}x_0 + \int_0^t e^{(t-s)A}Bu(s)\, ds\right) + \int_0^{T-t} e^{(T-t-s)A}Bu(s+t)\, ds \\ &= e^{TA}x_0 + \int_0^t e^{(T-s)A}Bu(s)\, ds + \int_t^T e^{(T-\delta)A}Bu(\delta)\, d\delta = y^{x_0,u}(T) = 0 \end{aligned}$$

So by definition, we have $T(x_1) \le T(x_0) - t$, i.e $T(x_0) \ge t + T(x_1) = t + T(y^{x_0,u}(t))$.

By step 1 and 2, we conclude that

$$T(x_0) = \inf_{u \in \mathcal{U}_{ad}} \{t + T(y^{x_0,u}(t))\} = \min_{u \in \mathcal{U}_{ad}} \{t + T(y^{x_0,u}(t))\}$$

Since $T(x_0) = t + T(y^{x_0,u}(t))$ where $u \in \mathcal{U}_{ad}$ such that $y^{x_0,u}(T(x_0)) = 0$. □

From above proof, we can deduce the principle for the optimality, i.e

Theorem 2.69 (Principle of Optimality)**:** Assume $T(x_0) = T$ with the optimal control $u \in \mathcal{U}_{ad}$, then for $t \in [0, T]$, setting $x_1 = y^{x_0,u}(t)$, then

$$T(x_1) = T - t$$

Proof. By dynamic programming principle we already prove above

$$T = T(x_0) = T(y^{x_0,u}(0)) \leq t + T(y^{x_0,u}(t)) = t + T(x_1) \Longrightarrow T(x_1) \geq T - t$$

Now doing similarly to the proof in dynamic programming principle, setting

$$v(s) = \begin{cases} u(s+t) & 0 \leq s \leq T - t \\ 0 & \text{otherwise} \end{cases}$$

then we have

$$T(y^{x_1,v}(T-t)) = 0 \Longrightarrow T(x_1) \leq T - t$$

Therefore $T(x_1) = T - t$. □

2.3.7 The Lipschitz and absolutely continuous properties of the trajectory

In this section, by using the formula (2.18) and the principle of optimality 2.69, we will establish the locally Lipschitz of $y^{x_0,u}(.)$, which is useful later. First we consider consider a bound for $\mathcal{R}(T)$

Proposition 2.70: Consider $||A|| > 0$, then

(i) For any $x_0 \in \partial\mathcal{R}(T)$, we have

$$||x_0|| \leq \frac{||B||}{||A||} e^{T||A||} \quad \Longleftrightarrow \quad \mathcal{R}(T) \subseteq B\left(0, ||B||.||A||^{-1} e^{T||A||}\right)$$

(ii) Furthermore, let's call $u \in \mathcal{U}_{ad}$ is the corresponding optimal control. Let $\mathbf{x}(.) = y^{x_0,u}(.)$ is the corresponding trajectory. Then for every $t \in [0, t]$, by principle of optimality 2.69, we know that

$$T(\mathbf{x}(t)) = T - t$$

so $\mathbf{x}(t) \in \partial\mathcal{R}(T - t)$ and therefore

$$||\mathbf{x}(t)|| \leq \frac{||B||}{||A||} e^{(T-t)||A||} \quad \Longleftrightarrow \quad \mathcal{R}(T-t) \subseteq B\left(0, ||B||.||A||^{-1} e^{(T-t)||A||}\right)$$

Proof.

(i) By using the formula (2.18), any $x_0 \in \mathcal{R}(T)$ has the form

$$x_0 = -\int_0^T e^{-sA} Bu(s)\, ds$$

for $u \in \mathcal{U}_{ad}$ is a control. Thus we have

$$\|x_0\| \leq \int_0^T e^{s\|A\|} \|B\|\, ds = \|B\| \int_0^T e^{s\|A\|}\, ds = \frac{\|B\|}{\|A\|} e^{T\|A\|}$$

(ii) It's just the consequence of (ii).

□

Theorem 2.71: Assume $T = T(x_0)$. There exists a constant $M = M(T) > 0$ such that

$$\|\mathbf{x}(t) - \mathbf{x}(0)\| \leq Mt \qquad \forall\, 0 \leq t \leq T$$

i.e, $\mathbf{x}(.)$ is Lipschitz at 0. We can choose $M = 2\|B\| e^{2T\|A\|}$.

Proof. In view of $\mathbf{x}(.) = y^{x_0,u}(.)$ is the optimal trajectory with optimal control $u \in \mathcal{U}_{ad}$, first using (2.60) to have the following estimate for any $t \in [0, T]$

$$e^{tA} - 1 = \int_0^t Ae^{sA}\, ds \Longrightarrow \|e^{tA} - 1\| \leq \int_0^t \|A\| e^{s\|A\|}\, ds \leq \|A\| e^{t\|A\|} t \tag{2.35}$$

Now using this fact and the formula

$$\mathbf{x}(t) - \mathbf{x}(0) = (e^{tA} - 1)x_0 + \int_0^t e^{(t-s)A} Bu(s)\, ds$$

we obtain the estimate (we have used proposition 2.70)

$$\begin{aligned}
\|\mathbf{x}(t) - \mathbf{x}(0)\| &\leq \|e^{tA} - 1\|.\|x_0\| + \left| \int_0^t e^{(t-s)A} Bu(s)\, ds \right| \\
&\leq \| \left(A\| e^{t\|A\|} t\right) \left(\frac{\|B\|}{\|A\|} e^{T\|A\|} \right) + \int_0^t e^{(t+s)\|A\|} \|B\|\, ds \\
&\leq \|B\| e^{(t+T)\|A\|} t + \|B\| e^{t\|A\|} \int_0^t e^{s\|A\|}\, ds \\
&\leq \|B\| e^{(t+T)\|A\|} t + \|B\| e^{2t\|A\|} t \\
&\leq 2\|B\| e^{(t+T)\|A\|} t \leq 2\|B\| e^{2T\|A\|} t
\end{aligned}$$

So we can choose $M = 2\|B\| e^{2T\|A\|}$ in the Lipschitz property at 0 of $\mathbf{x}(.)$. □

From the locally Lipschitz property at 0 of the trajectory, we can deduce that it's is Lipschitz in whole $[0, T]$.

Proposition 2.72: Assume $T = T(x_0)$. There exists a constant $M = M(T) > 0$ such that

$$\|\mathbf{x}(t) - \mathbf{x}(t')\| \leq M(t - t') \qquad \forall\, 0 \leq t, t' \leq T$$

i.e, $\mathbf{x}(.)$ is Lipschitz in $[0, T]$ with constant $M = \|B\|e^{2T\|A\|}$.

Proof. The proof is straight-forward, assume $0 \leq t < t' \leq T$, by principle of optimality 2.69, we have

$$T(\mathbf{x}(t)) = T - t$$

and if we consider the new control

$$v(s) = \begin{cases} u(s+t) & 0 \leq s \leq T - t \\ 0 & \text{otherwise} \end{cases}$$

then $T(y^{\mathbf{x}(t),v}(T - t)) = 0$, i.e v is the optimal control which steer $\mathbf{x}(t)$ to the origin in time $T - t$, thus $\mathbf{x}(t) \in \partial\mathcal{R}(T - t)$. Now also by proposition 2.70 we obtain

$$\|\mathbf{x}(t)\| \leq \frac{\|B\|}{\|A\|} e^{(T-t)\|A\|} \leq \frac{\|B\|}{\|A\|} e^{T\|A\|}$$

Consider the trajectory $\mathbf{z}(.) = y^{\mathbf{x}(t),v}(.)$, we have

$$\mathbf{z}(0) = \mathbf{x}(t) \qquad \text{and} \qquad \mathbf{z}(t' - t) = \mathbf{x}(t')$$

Thus, by using theorem 2.71 we obtain

$$\|\mathbf{x}(t') - \mathbf{x}(t)\| = \|\mathbf{z}(t' - t) - \mathbf{z}(0)\| \leq 2\|B\|e^{2T\|A\|} \cdot |t' - t|$$

Therefore we conclude that $\mathbf{x}(.)$ is Lipschitz in $[0, T]$ with constant $M = 2\|B\|e^{2T\|A\|}$. $\square$

As a corollary, we obtain the absolutely continuous of the trajectory, first recall that a function $f : [a, b] \longrightarrow \mathbb{R}^n$ is said to be absolutely continuous if for every $\varepsilon > 0$, there exists $\delta > 0$ such that for any $\{(a_k, b_k)\}_{k=1}^m$ is the collection of finite disjoint subintervals of $[a, b]$, then

$$\sum_{k=1}^{m} |b_k - a_k| < \delta \Longrightarrow \sum_{k=1}^{m} \|f(b_k) - f(a_k)\| < \varepsilon$$

If f is Lipschitz with constant C, then for any given $\varepsilon > 0$, we just choose $\delta = \frac{\varepsilon}{2C}$, then the absolutely continuous property of f will follow. Thus, the Lipschitz property implies absolutely continuous property of the trajectory.

Proposition 2.73: Given $x_0 \in \mathbb{R}^n$, $T(x_0) = T$ and $u \in \mathcal{U}_{ad}$ is a optimal control of x_0. Then the trajectory $\mathbf{x}(.) = y^{x_0,u}(.) : [0, T] \longrightarrow \mathbb{R}^n$ is absolutely continuous.

2.4 Definition of viscosity solutions

We consider the first order partial differential equation

$$F(x, u(x), \nabla u(x)) = 0 \tag{2.36}$$

where $F : \Omega \times \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}$ is a continuous function from an open set $\Omega \subset \mathbb{R}^n$

Definition 2.74: A function $u \in C(\Omega)$ is a **viscosity subsolution** of (2.36) if

$$F(x, u(x), p) \leq 0 \qquad \text{for every } x \in \Omega, p \in D^+u(x) \tag{2.37}$$

Similarly, $u \in C(\Omega)$ is **viscosity supersolution** of (2.36) if

$$F(x, u(x), p) \geq 0 \qquad \text{for every } x \in \Omega, p \in D^-u(x) \tag{2.38}$$

We say u is a viscosity solution of (2.36) if it is both a supersolution and a subsolution in the viscosity sense.

Note that our definition is well-define, since for $u \in C(\Omega)$, the set Ω^- and Ω^+ are both non-empty and dense in Ω. Furthermore, if $u \in C^1(\Omega)$ is a function satisfies (2.36), then by proposition 2.12, for any $x \in \Omega$ we have

$$D^+u(x) = D^-u(x) = \{\nabla u(x)\}$$

and clearly u is a viscosity solution of (2.36). Conversely, if u is a viscosity solution, then (2.36) must hold for any $x \in \Omega$ such that u is differentiable at x.

Now we define the terminology of viscosity solutions for the general Hamilton-Jacobi equations of the form

$$u_t + H(t, x, u, \nabla u) = 0 \qquad (t, x) \in (0, T) \times \Omega \tag{2.39}$$

where ∇u denotes the gradient of u which respect to x. From lemma 2.13 we have the following definition of viscosity solution for above equation

Definition 2.75: A continuous function $u : (0, T) \times \Omega \longrightarrow \mathbb{R}$ is a viscosity subsolution of (2.39) if for every C^1 function $\varphi = \varphi(t, x)$ such that $u - \varphi$ has a local maximum at (t, x), one has

$$\varphi_t(t, x) + H(t, x, u, \nabla \varphi) \leq 0$$

Similarly, a continuous function $u : (0, T) \times \Omega \longrightarrow \mathbb{R}$ is a viscosity supersolution of (2.39) if for every C^1 function $\varphi = \varphi(t, x)$ such that $u - \varphi$ has a local minimum at (t, x), one has

$$\varphi_t(t, x) + H(t, x, u, \nabla \varphi) \geq 0$$

3 Controllability

In this section, we want to study the basis controllability question: given the initial point x_0 and a "target" $S \subset \mathbb{R}^n$, is there a control which is steer the system to $S = \{0\}$ in finite time?

Definition 3.1: Consider the linear system (2.25), we call this system is

- **Small time controllable** on 0 if 0 is an interior point of $\mathcal{R}$
- **Fully controllable** if $\mathcal{R} = \mathbb{R}^n$.

To study the controllability of (2.25), we recall that $x_0 \in \mathcal{R}(t)$ iff exist $u(.) \in \mathcal{U}_{ad}$ such that

$$0 = e^{tA}x_0 + \int_0^t e^{(t-s)A}Bu(s)\, ds \Longleftrightarrow x_0 = -\int_0^t e^{-sA}Bu(s)\, ds$$

A simple example for the small time controllability. Let $n = 2, m = 1$ and $U = [-1, 1]$, consider the linear control system

$$\begin{cases} x'(t) & = bu(t) \qquad t > 0 \\ x(0) & = x_0 \end{cases}$$

where $b = (0, 1)^T$, we can easily compute that $\mathcal{R} = \{(x_1, x_2) : x_1 = 0\}$ so $\mathcal{R}$ cannot contain a neighborhood of the origin, i.e the control system is not small time controllable in 0.

Now we introduce some algebraic condition which ensure that $\mathcal{R}$ contains a small ball with the center being the origin. Let's start with a simple case, that is $B = b \in \mathbb{R}^n$.

3.1 Simple case $B = b \in \mathbb{R}^n$

Proposition 3.2: Assuming that $B = b \in \mathbb{R}^n$ and $U = [-1, 1]$, then

$$\mathcal{R} \subseteq H(A, b) = \text{span}\{b, Ab, \ldots, A^{n-1}b\}$$

Proof. Fixing $t > 0$, recall that

$$\mathcal{R}(t) = \left\{ x = -\int_0^t e^{-sA}bu(s)\, ds : u \in \mathcal{U}_{ad} \right\} \tag{3.1}$$

Let $P(\lambda) = \det(\lambda I_n - A)$ is the characteristic polynomial of A, by Hamilton–Cayley theorem we have $P(A) = 0$, i.e if we assume $P(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0$ then

$$A^n + a_{n-1}A^{n-1} + \ldots + a_1A + a_0I_n = 0 \Longleftrightarrow A^n = -a_{n-1}A^{n-1} - \ldots - a_1A - a_0I_n$$

So clearly A^k can be written as a linear combination of $\{I_n, A, \ldots, A^{n-1}\}$ for all $k \geq n$. Then because $e^{-sA}b = \sum_{k=0}^{\infty} \frac{(-s)^k A^k b}{k!}$ we conclude that $e^{-sA}b \in H(A, b) = \text{span}\{b, Ab, \ldots, A^{n-1}b\}$. From (3.1) we have $\mathcal{R}(t) \subseteq H(A, B)$ and complete the proof. $\square$

From this fact, we have

Proposition 3.3: For every $t > 0$, 0 is an interior point of $\mathcal{R}(t)$ in $H(A, b)$, i.e there exists $r(t) > 0$ such that $B(0, r(t)) \cap H(A, b) \subset \mathcal{R}(t) \subset \mathcal{R}$.

Proof. Assuming 0 isn't an interior point of $\mathcal{R}(t)$ in $H(A, b)$, then clearly it's must lie in the boundary of $\mathcal{R}(t)$. Recall that $\mathcal{R}(t)$ is compact and convex in $H(A, b)$, there exists a supporting hyperplane at 0, i.e there exists a unit vector $\xi \in H(A, b)$ such that

$$\langle \xi, y \rangle \leq 0 \qquad \forall\, y \in \mathcal{R}(t) \tag{3.2}$$

Now, let's consider $g(s) = \langle \xi, e^{-sA} b \rangle$ for all $s \in [0, +\infty)$. We can be easily seen that g is continuous from $\mathbb{R} \longrightarrow \mathbb{R}$, so if we the define the control $\overline{u} := -\text{sign}\Big(g(s)\Big)$ for $x \in [0, +\infty)$ then clearly $\overline{u} \in \mathcal{U}_{ad}$. Now from (3.2) we have

$$y = -\int_0^t e^{-sA} b \overline{u}(s)\, ds \in \mathcal{R}(t) \Longrightarrow \langle \xi, y \rangle = \int_0^t |g(s)|\, ds \leq 0$$

Since g is continuous, it implies that $g \equiv 0$ on $[0, t]$, hence $g^{(k)} \equiv 0$ on $[0, t]$ for all $k \in \mathbb{N}$. Therefore $\langle \xi, A^k b \rangle = 0$ for all $k = \overline{0, n-1}$, it's a contradiction to $\xi \in H(A, b)$. $\square$

By this proposition, if $\mathcal{R} = H(A, b)$ then the linear system (2.25) in case $B = b \in \mathbb{R}^n$ and $U = [-1, 1]$ is small time controllable, we lead to a result.

Theorem 3.4: Assuming that Re $\lambda \leq 0$ for each eigenvalue of A, then $\mathcal{R} = H(A, b)$.

Proof. Assume by a contradiction, we will introduce some steps, note that $\mathcal{R} \subseteq H(A, b)$ since proposition 3.2

1. If $\mathcal{R} \subsetneq H(A, b)$, then exists $z \in H(A, b) \backslash \mathcal{R}$. Since $\mathcal{R}$ is convex in $H(A, b) \subset \mathbb{R}^n$, by Hahn–Banach theorem, exist a hyperplane $[F = \mu]$ with $F : \mathbb{R}^n \longrightarrow \mathbb{R}$ separate $\mathcal{R}$ and $\{z\}$, i.e exists a vector $\xi \in H(A, b)$ and a real number μ such that

$$\langle \xi, y \rangle \leq \mu = F(z) \qquad \forall\, y \in \mathcal{R} \tag{3.3}$$

2. We will find some $y \in \mathcal{R}$ so that (3.3) is fails. Recall that $y \in \mathcal{R}$ iff exits a time $t > 0$ and a control $\alpha(.) \in \mathcal{U}_{ad}$ such that

$$y = -\int_0^t e^{-sA} b \alpha(s)\, ds \Longrightarrow \langle \xi, y \rangle = -\int_0^t \xi^T e^{-sA} b \alpha(s)\, ds$$

   Define $v(s) = \xi^T e^{-sA} b$ for all $s \in [0, \infty)$, then $\langle \xi, y \rangle = F(y) = -\int_0^t v(s)\alpha(s)\, ds$.

3. We assert that $v \not\equiv 0$. Suppose that $v \equiv 0$, then k–time differentiate the expression $v(s) = b^T e^{-sA} B$ with respect to s and set $s = 0$ to discover $\xi^T A^k b = 0$ for all $k \in \mathbb{N}$. It's a contradiction since $\xi \in H(A, b)$. So $v \not\equiv 0$.

4. Define $\alpha(s) = -\text{sign}(v(s))$ for all $s \in [0, \infty)$. It's obvious that $\alpha(.) \in \mathcal{A}$, let

$$x_0 = -\int_0^t e^{-sA} b\alpha(s)\, ds \Longrightarrow F(x_0) = \langle \xi, x_0 \rangle = -\int_0^t v(s)\alpha(s)\, ds = \int_0^t |v(s)|\, ds$$

Clearly $x_0 \in \mathcal{R}$, we will prove that x_0 make (3.3) to be fail.

5. To find $t > 0$ such that $F(x^0) > \mu$, we will prove that $\int_0^\infty |v(s)|\, ds = +\infty$. Assume it's not true, then the following function is well–define

$$\varphi(t) = \int_t^\infty v(s)\, ds \qquad \text{and} \qquad \varphi'(t) = -v(t) \quad \forall\, t > 0$$

Take $p(\lambda) = \det(A - \lambda I_n)$ to be the characteristic polynomial of A. Then according to the Caley–Hamilton theorem $p(A) = 0$. So we have

$$p\left(-\frac{d}{dt}\right) v(t) = p\left(-\frac{d}{dt}\right)\left(\xi^T e^{-tA} b\right) = \xi^T (p(A) e^{-tA}) b \equiv 0$$

So $v(t)$ is a solution of n^{th} order ODE $p\left(\frac{d}{dt}\right) v(t) = 0$. So $\varphi(t)$ is a solution of the $(n+1)^{\text{th}}$ order ODE

$$\left(-\frac{d}{dt}\right) p\left(-\frac{d}{dt}\right) \varphi(t) = 0$$

The polynomial $\lambda p(-\lambda) = 0$ has $n+1$ solutions $0, \lambda_1, \ldots, \lambda_n$, so $\mathcal{F}(t)$ has the form

$$\varphi(t) = \sum_{i=1}^n p_i(t) e^{-\lambda_i t} \tag{3.4}$$

where $\{p_i\}_{i=1}^n$ are appropriate polynomials. On the other hand, we have

$$|\varphi(t)| = \left| \int_t^\infty v(s)\, ds \right| \leq \int_t^\infty |v(s)|\, ds \longrightarrow 0 \quad \text{as } t \longrightarrow \infty$$

So $\varphi(t) \longrightarrow 0$ as $t \longrightarrow \infty$, but since Re $(\lambda_i) \leq 0$, we have a contradiction to the representation formula of (3.4). So $\int_0^\infty |v(s)|\, ds = +\infty$.

From this fact, exist a time t and x_0 such that $\langle \xi, x_0 \rangle = F(x_0) > \mu$, it's a contradiction to (3.3)!. So $\mathcal{R} = H(A, b)$ and we complete the proof. □

From above result, we can introduce a algebraic condition which ensures that the linear control system (2.25) is small time controllable. First define the *Controllability matrix* is

$$G(A, B) = [B, AB, \ldots A^{n-1} B]$$

which is a $n \times mn$ matrix.

So in case $B = b \in \mathbb{R}^n$ and $U = [-1, 1]$, we have $G(A, b)$ is a $n \times n$ matrix and we have the following result

Theorem 3.5: Let $B = b \in \mathbb{R}^n$ and $U = [-1,1]$, then the linear system (2.25) is small time controllable if and only if

$$\text{rank } G(A,b) = n$$

Proof. Assume that the linear system (2.25) is small time controllable, i.e there exists $\delta > 0$ such that $B(0,r) \subset \mathcal{R} \subseteq H(A,b) = \text{span}\{b, Ab, \ldots, A^{n-1}b\}$ by proposition 3.2. Then it follow that rank $G(A,b) = n$. For the converse, if rank $G(A,b) = n$ then $H(A,b) = \mathbb{R}^n$ and so clearly 0 is an interior point of $\mathcal{R}$ by proposition 3.3, i.e the linear system (2.25) is small time controllable. □

Similarly, from proposition 3.4 we conclude that

Theorem 3.6: Let $B = b \in \mathbb{R}^n$ and $U = [-1,1]$. Assume that Re $\lambda \leq 0$ for each eigenvalue of A and rank $G(A,b) = n$. Then the linear system (2.25) is fully controllable.

Proof. If rank $G(A,b) = n$ then $H(A,b) = \mathbb{R}^n$, and from proposition 3.4 we conclude that $\mathcal{R} = H(A,b) = \mathbb{R}^n$. □

3.2 General case $B \in \mathbb{M}^{n\times m}(\mathbb{R})$

In general case $B \in \mathbb{M}^{n\times m}(\mathbb{R})$, we have a similar result of proposition 3.2 and 3.3

Proposition 3.7: For every $t > 0$, we have $\mathcal{R}(t) \subset H(A,B)$ and there exists $r_t > 0$ s.t

$$B(0,r_t) \cap H(A,B) \subset \mathcal{R}(t)$$

Proof. Write $B = [b_1, b_2, \ldots, b_m]$ where $b_i \in \mathbb{R}^n$ is columns of B for $i = \overline{1,m}$, note that control $u : [0,\infty) \longrightarrow [-1,1]^m$ can be representation of the form $u = (u_1, u_2, \ldots, u_m)^T$ where $u_i : [0,\infty) \longrightarrow [-1,1]$. One has that

$$-\int_0^t e^{-sA} Bu(s)\, ds = -\sum_{i=1}^m \int_0^t e^{-sA} b_i u_i(s)\, ds$$

and hence $\mathcal{R}(t) = \mathcal{R}(t)_1 + \ldots + \mathcal{R}(t)_m$ where $\mathcal{R}(t)_i$ is the reachable set in time t of the linear control system (2.25) with $B = b_i$ for all $i = \overline{1,m}$. On the other hand, note that clearly $H(A,B) = H(A,b_1) + \ldots + H(A,b_m)$. Now from proposition 3.2, we have that $\mathcal{R}(t)_i \subseteq H(A,b_i)$ for all $i = \overline{1,m}$, so clearly $\mathcal{R} \subseteq H(A,b)$. For the rest, from proposition 3.3 for each $i \in \overline{1,m}$ there exists $r_i > 0$ such that

$$B(0,r_i) \cap H(A,b_i) \subset \mathcal{R}(t)_i \subseteq \mathcal{R}(t)$$

Let $r = \min\{r_i : i \in \overline{1,m}\}$, then clearly $B(0,r) \cap H(A,B) \subset \mathcal{R}(t)$, the proof is complete. □

Now we state some condition for small time controllability and fully controllability for the linear control system (2.25) in case $B \in M^{n\times m}(\mathbb{R})$.

Theorem 3.8: The linear control system (2.25) is small time controllable if and only if rank $G(A,B) = n$.

Proof. According to the proof of theorem 3.7, we have that $\mathcal{R} = \mathcal{R}_1 + \ldots + \mathcal{R}_m$ where $\mathcal{R}_i$ is the reachable set of the linear control system (2.25) with $B = b_i$ for all $i \in \overline{1,m}$.
Now if rank $G(A,B) = n$, then $H(A,B) = \mathbb{R}^n$, from proposition 3.7 we have that there exists $r > 0$ such that $B(0,r) = B(0,r) \cap H(A,B) \subset \mathcal{R}$, i.e the linear system (2.25) is small time controllable. For the converse, if the linear system (2.25) is small time controllable, i.e there exists $r > 0$ such that $B(0,r) \subset \mathcal{R}$, we will prove that rank $G(A,B) = n$. Assume by a contradiction, one has $\dim H(A,B) < n$, from proposition 3.7 we have $\mathcal{R} \subseteq H(A,B)$ so $\dim \mathcal{R} \leq n-1$, which is cannot contain any ball at 0, it's a contradiction. □

We conclude this section with a result about fully controllable condition of the linear system (2.25) in general case.

Theorem 3.9: Assume that rank $G(A,B) = n$ and Re $\lambda \leq 0$ for each eigenvalue of A. Then the linear control system 2.25 is fully controllable, i.e $\mathcal{R} = \mathbb{R}^n$.

Proof. First note that if rank $G(A,B) = n$ then $H(A,B) = \mathbb{R}^n$, we will claim that $\mathcal{R} = H(A,B)$ under this assumption that Re $\lambda \leq 0$ for each eigenvalue of A. From 3.4 for all $i \in \overline{1,m}$ we have $\mathcal{R}_i = H(A, b_i)$ where $\mathcal{R}_i$ is the reachable set of the linear control system (2.25) with $B = b_i$. So clearly

$$\mathcal{R} = \mathcal{R}_1 + \ldots + \mathcal{R}_m = H(A, b_1) + \ldots + H(A, b_m) = H(A,B) = \mathbb{R}^n$$

and the proof is complete. □

3.3 The continuity of the minimum time function

Now we study the continuity property of T.

Theorem 3.10: Let $T_{\mathcal{R}} : \mathcal{R} \longrightarrow [0, \infty)$ be denoted by $T_{\mathcal{R}}(x) = T(x)$ for all $x \in \mathcal{R}$, then $x \longmapsto T_{\mathcal{R}}(x)$ is continuous from $\mathcal{R} \longrightarrow \mathbb{R}$.

Proof. First we claim that $T_{\mathcal{R}}$ is continuous at 0. Indeed, let $\{x_k\} \subset \mathcal{R}$ be a sequence converging to 0, recall that for every $t > 0$ there exists $r(t) > 0$ such that

$$B(0, r(t)) \cap H(A,B) \subset \mathcal{R}(t)$$

Let $\{t_k\}$ be the sequence of real numbers converging to 0 such that $x_k \in R(t_k)$, it implies that $0 \leq T_{\mathcal{R}}(x_k) \leq t_k$ for all $k \in \mathbb{N}$. So since $t_k \longrightarrow 0$, hence $T_{\mathcal{R}}$ is continuous at 0.

Now assume $x \in \mathcal{R}\backslash\{0\}$, for $z \in \mathcal{R}$ such that $T_{\mathcal{R}}(z) \geq T_{\mathcal{R}}(x)$. Let $u_x \in \mathcal{U}_{ad}$ such that $y^{x,u_x}(T(x)) = 0$, i.e

$$y^{x,u_x}(T(x)) = e^{T(x)A}x + \int_0^{T(x)} e^{(T-s)A}Bu_x(s)\,ds = 0 \Longleftrightarrow \int_0^{T(x)} e^{(T-s)A}Bu_x(s)\,ds = -e^{T(x)A}x$$

Consider the trajectory $y^{z,u_x}(.)$ and define $\overline{z} = y^{z,u_x}(T(x))$ we can see that

$$\overline{z} = y^{z,u_x}(T(x)) = e^{T(x)A}z + \int_0^{T(x)} e^{(T-s)A}Bu_x(s)\,ds = e^{T(x)A}(z-x)$$

So we have $\|\overline{z}\| \leq e^{T(x)\|A\|}\|z-x\| = \delta_z$. By dynamics programming principle we have

$$T(z) \leq T(x) + T(\overline{z}) \Longrightarrow T(z) - T(x) \leq T(\overline{z}) \leq \sup_{z^* \in B'(0,\delta_z)} T(z^*) \tag{3.5}$$

Similarly, if $z \in \mathbb{R}$ such that $T(z) \leq T(x)$, let $u_z \in \mathcal{U}_{ad}$ such that $y^{z,u_z}(T(z)) = 0$, i.e

$$y^{z,u_z}(T(z)) = e^{T(z)A}z + \int_0^{T(z)} e^{(T-s)A}Bu_z(s)\,ds = 0 \Longleftrightarrow \int_0^{T(z)} e^{(T-s)A}Bu_z(s)\,ds = -e^{T(z)A}z$$

Consider the trajectory $y^{x,u_z}(.)$ and define $\overline{x} = y^{x,u_z}(T(z))$ we can see that

$$\overline{x} = y^{x,u_z}(T(z)) = e^{T(z)A}x + \int_0^{T(z)} e^{(T-s)A}Bu_z(s)\,ds = e^{T(x)A}(x-z)$$

So we have $\|\overline{x}\| \leq e^{T(x)\|A\|}\|z-x\| \leq e^{T(x)\|A\|}\|z-x\| = \delta_z$. By dynamics programming principle we have

$$T(x) \leq T(z) + T(\overline{x}) \Longrightarrow T(x) - T(z) \leq T(\overline{x}) \leq \sup_{z^* \in B'(0,\delta_z)} T(z^*) \tag{3.6}$$

From (3.5) and (3.6) we conclude that for all $z \in \mathcal{R}$ we have

$$|T_R(x) - T_R(z)| \leq \sup_{z^* \in B'(0,\delta_z)} T(z^*)$$

where $\delta_z = e^{T(x)\|A\|}\|z-x\|$, since $T_{\mathcal{R}}$ is continuous at 0, we have $T_{\mathcal{R}}$ is continuous at x. $\square$

Finally, if rank $G(A,B) = n$ and Re$\lambda \leq 0$ for each eigenvalue of A, we have $\mathbb{R}^n = H(A,B) = \mathcal{R}$, so the time minimum function $T_{\mathcal{R}}$ will continuous on $\mathbb{R}^n$.

Proposition 3.11: Let $T : \mathcal{R} \longrightarrow \mathbb{R}$ is the continuous minimum time function for the control system (2.25). If this system satisfies Re $\lambda \leq 0$ for each eigenvalue λ of A, and the Rank-Kalman's condition, that is

$$\text{rank } G(A,B) = n$$

then $\mathcal{R} = \mathbb{R}^n$, thus $T : \mathbb{R}^n \longrightarrow \mathbb{R}$ is continuous.

4 Optimal controls

4.1 Bang–Bang principle for linear control systems

The control $u \in \mathcal{U}_{ad}$ is called **bang–bang** if for each time $t \geq 0$, we have $|u_i(t)| = 1$ for all $i = \overline{1,m}$, where

$$u(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_m(t) \end{pmatrix}$$

Theorem 4.1 (Bang–bang principle): Given any $\overline{x} \in \mathcal{R}$, then there exists a bang–bang control u which steer $\overline{x}$ to the origin.

Recalling that

$$\mathcal{R} = \mathcal{R}_1 + \mathcal{R}_2 + \ldots + \mathcal{R}_m$$

where $\mathcal{R}_i$ is the reachable set of the system (2.25) with $B = b_i$ for all $i = \overline{1,m}$. So to prove Bang–bang principle, we only need to show in the case of $B = b \in \mathbb{R}^n$ and $U = [-1,1]$. In this case, recall that for every $t > 0$ we have

$$\mathcal{R}(t) \subseteq H(A,b) = \text{span}\,\{b, Ab, \ldots, A^{n-1}b\}$$

Proposition 4.2: For every $t > 0$, consider the reachable set $\mathcal{R}(t)$. Let x be on the boundary of $R(t)$ in the space $H(A,b)$. Then exists an bang–bang control which steers x to the origin in time t.

Proof. Clearly $R(t)$ is convex in $H(A,b)$, thus for every $x \in \partial\mathcal{R}(t)$, exists a unit vector $\xi_x \in H(A,b)$ suth that

$$(\xi_x)^T \cdot (y - x) \leq 0 \qquad \forall\, y \in \mathcal{R}(t) \tag{4.1}$$

Note that $x = -\int_0^t e^{-sA} b u(s)\, ds$, where $u : [0,t) \longrightarrow U$ is measurable. Let's consider

$$\begin{aligned} g : \mathbb{R} &\longrightarrow \mathbb{R} \\ s &\longmapsto \xi_x^T e^{-sA} b \end{aligned}$$

It's easy to check that g is a continuous mapping. Further more, it's infinitely differentiable, so we claim that the set Ker $g = \{s \in \mathbb{R} : g(s)\} = 0$ is finite. Indeed, if this set is infinite, we can choose a sequence $\{s_n\} \subset$ Ker g such that $s_n < s_{n+1}$ for any $n \in \mathbb{N}$. From mean–value theorem, we have $s_{k,k+1} \in (s_k, s_k + 1)$ such that

$$g' s_{12} = g'(s_{23}) = g'(s_{34}) = \ldots = g'(s_{k,k+1}) = \ldots \qquad \forall\, k \in \mathbb{N}$$

So similarly, we obtain a increase sequence t_k such that $g^{(n-1)}(t_k) = 0$ for all $k \in \mathbb{N}$, i.e $\xi_x^T \cdot A^k b = 0$ for all $k = 0, 1, 2 \ldots, n-1$. It's a contradiction since $\xi_x \in H(A,b)$ so cannot orthogonal to any column of matrix $[b, Ab, A^2 b, \ldots, A^{n-1} b]$.

Define

$$\overline{u}(s) = -\text{sign}\, g(s) \qquad \forall\, s \in [0, +\infty) \tag{4.2}$$

Since g is continuous, it's clearly $\overline{u} \in \mathcal{U}_{ad}$, hence $\overline{y} = -\int_0^t e^{-sA} b\overline{u}(s)\, ds \in \mathcal{R}(t)$. Since (4.1)

$$\begin{aligned}\xi_x^T \cdot (\overline{y} - x) \leq 0 &\Longleftrightarrow \int_0^t \xi_x^T \Big[-e^{-sA} b\overline{u}(s) + e^{-sA} bu(s) \Big]\, ds \leq 0 \\ &\Longleftrightarrow \int_0^t |g(s)|.\Big[1 + \operatorname{sign}(g(s))u(s)\Big]\, ds \leq 0\end{aligned}$$

Clearly $1 + \operatorname{sign}(g(s))u(s) \geq 0$ since $u(x) \in [-1, 1]$, so we obtain

$$|g(s)|.\Big[1 + \operatorname{sign}(g(s))u(s)\Big] = 0 \quad \text{for a.e } s \in [0, t]$$

Since Ker g is finite, we also have

$$1 + \operatorname{sign}(g(s))u(s) = 0 \quad \text{for a.e } s \in [0, t] \Longleftrightarrow u(s) = \overline{u}(s) \quad \text{for a.e } s \in [0, t]$$

Therefore, x is also steered to the origin by a bang–bang control $\overline{u}$. □

Note that we also obtain that any control steer x to the origin in this case is a.e bang–bang. Now, we assert a property of $\mathcal{R}(t)$ base on above proposition.

Proposition 4.3: For every $t > 0$, the reachable set $\mathcal{R}(t)$ is strictly convex in space $H(A, b)$

Proof. Let $x_0 \neq y_0$ in $\mathcal{R}(t)$ which respect to control u, v, i.e

$$x_0 = -\int_0^t e^{-sA} bu(s)\, ds \qquad\qquad y_0 = -\int_0^t e^{-sA} bv(s)\, ds \tag{4.3}$$

Given $\lambda \in (0, 1)$, assume that there exist a point $z = \lambda x_0 + (1-t)\lambda y_0 \in \partial\mathcal{R}(t)$, we will lead to a contradiction. Indeed, z is respect to the control $\alpha = \lambda u + (1-\lambda)v$, since $z \in \partial\mathcal{R}(t)$, from proposition (4.2) we obtain that α is a bang–bang control, i.e for a.e $s \in [0, t]$ we have $|\alpha(s)| = 1$. Note that $\max\{|u(s)|, |v(s)|\} \leq 1$ for all $s \in [0, 1]$, so

$$|\alpha(s)| = |\lambda u(s) + (1-\lambda)v(s)| \leq \lambda |u(s)| + (1-\lambda)|v(s)| \leq 1$$

hence we also have $|u(s)| = |v(s)| = 1$ for a.e $s \in [0, t]$, i.e u, v are also bang–bang controls. Now we also have $u(s) = v(s)$ for a.e $s \in [0, t]$ because

$$\begin{aligned}\alpha(s) = 1 &\Longrightarrow 1 = \lambda u(s) + (1-\lambda)v(s) \leq \lambda + (1-\lambda) = 1 \Longrightarrow u(s) = v(s) = 1 \\ \alpha(s) = -1 &\Longrightarrow -1 = \lambda u(s) + (1-\lambda)v(s) \geq -\lambda - (1-\lambda) = -1 \Longrightarrow -u(s) = -v(s) = -1\end{aligned}$$

So from (4.3) we obtain $x_0 = y_0$, it's contradiction. Hence $R(t)$ is stritly convex in $H(A, b)$. □

We can easily prove proposition 4.3 by using the property of the strictly convex set, by proposition 2.19, indeed

Proof based on proposition 2.19. Let $x \in \partial\mathcal{R}(t)$ arbitrary, then from proposition 4.2, there exists a bang bang control $u(.)$ steer x to the origin in time t, i.e

$$x = -\int_0^t e^{sA} bu(s)\, ds$$

Consider a supporting hyperplane of $\mathcal{R}(t)$ at x, given by the unit normal vector ξ. All we need is prove that

$$\langle \xi, y\rangle < \langle \xi, x\rangle \qquad \forall\ y \in \mathcal{R}(t)$$

Assume by contradiction that there exists $y \in \mathcal{R}(t)$ such that $\langle \xi, x\rangle = \langle \xi, y\rangle$. Note that also from proposition 4.2, we have $u(s) = -\text{sign}\ g(s)$ a.e in $[0, t]$, where

$$g(s) = \xi^T \cdot e^{sA} b$$

So if $\langle \xi, x\rangle = \langle \xi, y\rangle$, assume y is respect to the control v, we must have

$$-\int_0^t g(s)u(s)\, ds = -\int_0^t g(s)v(s)\, ds = \int_0^t |g(s)|\, ds \Longrightarrow \int_0^t |g(s)|\Big[1+\text{sign}(g(s))v(s)\Big]\, ds = 0$$

Since $\text{sign}(g(s))v(s) + 1 \geq 0$ for all $s \in [0, t]$, we must have

$$\text{sign}(g(s))v(s) = -1 \quad \text{a.e in } [0, t] \Longrightarrow v(s) = -\text{sign}(g(s)) = u(s) \quad \text{a.e in } [0, t]$$

And clearly from this we have $x = y$, it's a contradiction. □

Finally, from propositions (4.2) we obtain a general result in the case $x \in \mathcal{R}$.

Theorem 4.4: Let $B = b \in \mathbb{R}^n$ and $U = [-1, 1]$. Then for every point $x \in \mathcal{R} = \bigcup_{t\geq 0} \mathcal{R}(t)$, there exists a bang–bang control which steers x to the origin.

Proof. We claim that for every $\lambda > 0$, there exists a number $\varepsilon > 0$ such that

$$\mathcal{R}(t) \subseteq (1+\lambda)\mathcal{R}(t-\varepsilon) \tag{4.4}$$

We will introduce some steps.

1. For $a, b > 0$, we have $\mathcal{R}(a+b) \subseteq \mathcal{R}(a) + e^{-aA}\mathcal{R}(b)$.
   Indeed, let $x \in \mathcal{R}(a+b)$, then exists a control $u \in \mathcal{U}_{ad}$ such that

$$\begin{aligned} x = -\int_0^{a+b} e^{-sA} Bu(s)\, ds &= -\int_0^a e^{-sA} Bu(s)\, ds - \int_a^{a+b} e^{-sA} Bu(s)\, ds \\ &= -\int_0^a e^{-sA} Bu(s)\, ds - e^{-aA}\int_0^b e^{-\delta A} Bu(a+\delta)\, d\delta \end{aligned}$$

   Define $v(s) := u(a+s)$ then clearly $v \in \mathcal{U}_{ad}$ and this implies $x \in \mathcal{R}(a) + e^{-aA}\mathcal{R}(b)$ since

$$x = -\int_0^a e^{-sA} Bu(s)\, ds - e^{-aA}\int_0^b e^{-sA} Bv(s)\, ds \in \mathcal{R}(a) + e^{-aA}\mathcal{R}(b)$$

2. For any $t > 0$, 0 is an interior point of $\mathcal{R}(t)$ in $H(A,b)$, i.e theres exist $r_t > 0$ such that $B(0,r_t) \cap H(A,b) \subset \mathcal{R}(t)$.

3. For $\lambda > 0$, there exists $\varepsilon > 0$ such that $\mathcal{R}(\varepsilon) \subseteq \lambda\mathcal{R}(t-\varepsilon)$.
Indeed, for any $x \in \mathcal{R}(\varepsilon)$ where $\varepsilon < 1$ will be choosing later, we have

$$\|x\| = \left\| \int_0^\varepsilon e^{-sA} Bu(s)\, ds \right\| \leq \int_0^\varepsilon \left\| e^{-sA} Bu(s) \right\| \, ds \leq \varepsilon e^{\varepsilon \|A\|} \|B\| \leq (e^{\|A\|} \|B\|)\varepsilon$$

So we have $\mathcal{R}(\varepsilon) \subseteq B(0, M\varepsilon)$ where $M = e^{\|A\|}\|B\|$. Now from step [2.], exists $r > 0$ such that $B(0,r) \cap H(A,b) \subset \mathcal{R}\left(\frac{t}{2}\right)$. Finally, choose $\varepsilon > 0$ small such that

$$\varepsilon < \min\left\{\frac{t}{2}, \frac{r\lambda}{M}\right\} \Longrightarrow \frac{t}{2} < t - \varepsilon, \frac{M\varepsilon}{\lambda} < r$$

Then we have $B(0, M\varepsilon) \subset \lambda B\left(0, \frac{M\varepsilon}{\lambda}\right)$, so

$$R(\varepsilon) = \mathcal{R}(\varepsilon) \cap H(A,b) \subseteq B(0,M\varepsilon) \cap H(A,b) \subset \lambda B\left(0, \frac{M\varepsilon}{\lambda}\right) \cap H(A,b) \subset \lambda\mathcal{R}\left(\frac{t}{2}\right) \subset \lambda\mathcal{R}(t-\varepsilon)$$

4. For any $x \in \mathcal{R}$, we have $e^{-\varepsilon A}x \in \mathcal{R}$.
Indeed, if $x \in \mathcal{R}(t)$, we have

$$e^{-\varepsilon A}x = -\int_0^t e^{-(\varepsilon+s)A} Bu(s)\, ds = \int_\varepsilon^{t+\varepsilon} e^{-\delta A} Bu(\delta - \varepsilon)\, d\delta$$

Define a new control

$$v(s) := \begin{cases} u(s-\varepsilon) & s \in [\varepsilon, t+\varepsilon] \\ 0 & s \in [0, \varepsilon) \end{cases}$$

then it's clear that

$$e^{-\varepsilon A}x = -\int_0^{t+\varepsilon} e^{-sA} Bv(s)\, ds \in \mathcal{R}(t+\varepsilon) \subset \mathcal{R}$$

5. From [4.], it's follow that $(e^{-\varepsilon A} - 1)x \in H(A,b)$ since $H(A,b) = \text{span}\,\{b, Ab, A^2b, \ldots, A^{n-1}b\}$.

6. For $\lambda > 0$, there exists $\varepsilon > 0$ such that $e^{-\varepsilon A}\mathcal{R}(t-\varepsilon) \subset (1+\lambda)\mathcal{R}(t-\varepsilon)$.
Indeed, let $x \in \mathcal{R}(t-\varepsilon)$, consider

$$e^{-\varepsilon A}x = \left(e^{-\varepsilon A} - 1\right)x + x$$

Let $r > 0$ such that $B(0,r) \cap H(A,b) \subset \mathcal{R}\left(\frac{t}{2}\right)$. On the other hand, we have $\mathcal{R}(t-\varepsilon) \subset B(0, M(t-\varepsilon)) \subset B(0, Mt)$, so

$$\sup_{x \in \mathcal{R}(t-\varepsilon)} \left\| \left(e^{-\varepsilon A} - 1\right)x \right\| \longrightarrow 0 \qquad \text{as} \qquad \varepsilon \longrightarrow 0$$

so we can choose $\varepsilon > 0$ so small such that $(e^{-\varepsilon A} - 1)x \in B(0, \lambda r)$ for all $x \in \mathcal{R}(t-\varepsilon)$, then

$$(e^{-\varepsilon A} - 1)x \in B(0, \lambda r) \cap H(A, b) \subseteq \lambda B(0, r) \cap H(A, b) \subseteq \lambda \mathcal{R}\left(\frac{t}{2}\right) \subset \lambda \mathcal{R}(t-\varepsilon)$$

So we have

$$e^{-\varepsilon A}x = \left(e^{-\varepsilon A} - 1\right)x + x \subset \lambda \mathcal{R}(t-\varepsilon) + \mathcal{R}(t-\varepsilon)$$

Assume that $e^{-\varepsilon A}x = \lambda a + b$ where $a, b \in \mathcal{R}(t-\varepsilon)$, then

$$e^{-\varepsilon A}x = (\lambda+1)\left[\frac{\lambda}{\lambda+1}a + \frac{1}{\lambda+1}b\right] \in (\lambda+1)\mathcal{R}(t-\varepsilon)$$

since $\mathcal{R}(t-\varepsilon)$ is convex. So we conclude that $e^{-\varepsilon A}\mathcal{R}(t-\varepsilon) \subset (\lambda+1)\mathcal{R}(t-\varepsilon)$.

Finally, we prove (4.4), indeed, let $a = \varepsilon$ and $b = t - \varepsilon$ in step [1.], we have

$$\mathcal{R}(t) \subseteq \mathcal{R}(\varepsilon) + e^{-\varepsilon A}\mathcal{R}(t-\varepsilon)$$

Now from step [3.] and [6.] there exists $\varepsilon > 0$ such that

$$\mathcal{R}(\varepsilon) \subseteq \frac{\lambda}{2}\mathcal{R}(t-\varepsilon) \qquad \text{and} \qquad e^{-\varepsilon A}\mathcal{R}(t-\varepsilon) \subseteq \left(\frac{\lambda}{2} + 1\right)\mathcal{R}(t-\varepsilon)$$

So we have

$$\mathcal{R}(t) \subseteq \mathcal{R}(\varepsilon) + e^{-\varepsilon A}\mathcal{R}(t-\varepsilon) \subseteq \frac{\lambda}{2}\mathcal{R}(t-\varepsilon) + \left(\frac{\lambda}{2} + 1\right)\mathcal{R}(t-\varepsilon) \subset (\lambda+1)\mathcal{R}(t-\varepsilon)$$

since $\mathcal{R}(t-\varepsilon)$ is convex. We complete the proof of (4.4).
Now, let $0 \neq x \in \mathcal{R} = \bigcup_{t\geq 0} \mathcal{R}(t)$, we will prove that there exists s such that $x \in \partial\mathcal{R}(s)$, and from 4.2 we have a bang–bang control which steers x to the origin. To see this, let

$$t = \inf\{\delta > 0 : x \in \mathcal{R}(\delta)\} \tag{4.5}$$

It's easy to see that $x \in \mathcal{R}(t)$,(theorem 2.29). Now, we claim that $x \in \partial\mathcal{R}(t)$, indeed, assume that there exists $r > 0$ such that $B(x, r) \subset \mathcal{R}(t)$, define $\lambda := \frac{r}{2\|x\|}$, we have

$$(\lambda+1)x \in B(x, r) \subset \mathcal{R}(t)$$

From (4.4), there exists $\varepsilon > 0$ such that

$$(\lambda+1)x \in \mathcal{R}(t) \subseteq (\lambda+1)\mathcal{R}(t-\varepsilon) \Longrightarrow x \in \mathcal{R}(t-\varepsilon)$$

It's a contradiction to (4.5), so we conclude that $x \in \partial\mathcal{R}(t)$. □

From this, we have a simple but very useful fact

Remark 4.5: We have conclude that if $x \in \mathcal{R}$, then if we set $t = \inf\{\delta > 0 : x \in \mathcal{R}(\delta)\}$ then $x \in \partial\mathcal{R}(t)$, so by theorem 4.2, there exists a bang–bang control u^* which steers x to 0 in time t. Clearly the bang–bang control u^* is an **optimal control**. Since $t = \inf\{\delta > 0 : x \in \mathcal{R}(\delta)\}$. So we obtain

Theorem 4.6: For every point $x \in \mathcal{R}$, there exists a bang–bang optimal control u^* which steers x to 0.

4.2 The maximum principle for linear control systems

In this section, we introduce a useful principle, which is the consequence of Bang-bang principle. From which, we can study easily about the switching time of the optimal controls.

Now recall theorem 4.6, we know that for every $x \in \mathcal{R}$, there exists a bang–bang optimal control u^* which steer x to the origin. Further more, we have

Theorem 4.7 (The Pontryagin maximum principle for linear control systems)**:** If u^* is a bang–bang optimal control which steer x to the origin in time t, there exists a nonzero vector θ such that

$$\langle \theta, e^{-sA}Bu^*(s)\rangle = \max_{\lambda\in U}\langle \theta, e^{-sA}B\lambda\rangle \tag{4.6}$$

for each time $s \in [0,t]$. Recall that $U = [-1,1]^m$ in this case.

Proof. We introduce some steps

1. A point $y \in \mathcal{R}(t)$, respect to the control u if and only if

$$y = -\int_0^t e^{-sA}Bu(s)\,ds \tag{4.7}$$

2. We know that $x \in \partial\mathcal{R}(t)$, and $\mathcal{R}(t)$ is strictly convex in $H(A,B)$, by supporting hyperplane theorem for strictly convex sets, there exists a unit vector $\xi \in H(A,B)$ such that

$$\langle \xi, y\rangle < \langle \xi, x\rangle \qquad \forall\, y \in \mathcal{R}(t)\backslash\{x\}$$

   Now from (4.7), we have

$$-\int_0^t \xi^T e^{-sA}Bu(s)\,ds < -\int_0^t \xi^T e^{-sA}Bu^*(s)\,ds \qquad \forall\, u \in \mathcal{U}_{ad}$$

   Setting $\theta = -\xi$, we have

$$\int_0^t \theta^T e^{-sA}B\Big(u^*(s) - u(s)\Big)\,ds > 0 \qquad \forall\, u \in \mathcal{U}_{ad} \tag{4.8}$$

3. We claim that (4.6) is true, i.e

$$\theta^T e^{-sA}Bu^*(s) = \max_{\lambda\in U}\left\{\theta^T e^{-sA}B\lambda\right\} \qquad \text{for a.e} \qquad s\in[0,t]$$

   Assuming by contradiction, there would exist a subset $E \subset [0,t]$ of positive measure such that

$$\theta^T e^{-sA}Bu^*(s) < \max_{\lambda\in U}\left\{\theta^T e^{-sA}B\lambda\right\} \qquad \forall\, s\in E \tag{4.9}$$

   For any $s \in [0,t]$, the mapping $g_s : U \longrightarrow \mathbb{R}$ where $g_s(\lambda) = \theta^T e^{-sA}B\lambda$ is continuous on a compact set $U = [-1,1]^m$, so it must achieve a maximum, denote by α_s.
   We want to construct a measurable control function $v : E \longrightarrow U$ such that

$$g_s \cdot v(s) = \alpha_s \qquad \text{for a.e } s \in E$$

Also note that for any $s \in [0,t]$, the mapping $g : s \longrightarrow g_s$ from $[0,t] \longrightarrow \mathbb{R}^m$ is also continuous and furthermore, it's in $C^\infty(0,t)$. And so similarly to the proof of theorem 4.2, we can see that the set Ker g is finite since $\theta \in H(A,B)$. In consider a simple case at first

- Case $m=1$, i.e $U=[-1,1]$, then clearly

$$g_s(\lambda) \leq |g_s| \quad \Longleftrightarrow \quad \alpha_s = |g_s|$$

since it's archived at $\lambda = \text{sign } g_s$. Naturally, we can choose $v(s) = \text{sign } g_s$ form $E \longrightarrow U = [-1,1]$. Clearly v is measurable.

- Case $U=[-1,1]^m$, we can assume that g_s is presented by the vector

$$g_s = (a_1(s), a_2(s), \ldots, a_m(s))^T$$

where $s \longmapsto a_i(s)$ are continuous mappings for $i = \overline{1,m}$. Then for any $\lambda = (\lambda_1, \ldots, \lambda_m) \in U$, we have

$$\theta^T e^{-sA} B\lambda = g_s(\lambda) = \begin{pmatrix} a_1(s) \\ \vdots \\ a_m(s) \end{pmatrix} \cdot (\lambda_1, \ldots, \lambda_m) = \sum_{i=1}^m a_i(s)\lambda_i \leq \sum_{i=1}^m |a_i(s)|$$

So $\alpha_s = \sum_{i=1}^m |a_i(s)|$ since it's archived at $(\lambda_1, \ldots, \lambda_m) = (\text{sign } a_1(s), \ldots, \text{sign } a_m(s))$. So clearly $v(s) = (\text{sign } a_1(s), \ldots, \text{sign } a_m(s))$ is measurable.

Now we design a new control $\overline{u}$ as follows

$$\overline{u}(s) = \begin{cases} u^*(s) & s \notin E \\ v(s) & s \in E \end{cases}$$

Clearly $\overline{u}$ is measurable since v and E is measurable. Now from (4.8) and (4.9), we have

$$\int_0^t \theta^t e^{-sA} B\Big(u^*(s) - \overline{u}(s)\Big)\, ds = \int_E \theta^t e^{-sA} B\Big(u^*(s) - \overline{u}(s)\Big)\, ds \geq 0$$

but for a.e $s \in E$ we also have

$$\theta^t e^{-sA} B u^*(s) < \max_{\lambda \in U}\{\theta^t e^{-sA} B\lambda\} = \theta^t e^{-sA} B\overline{u}(s) \Longrightarrow \int_E \theta^t e^{-sA} B\Big(u^*(s) - \overline{u}(s)\Big)\, ds \leq 0$$

It's a contradiction since $\mu(E) > 0$.

Finally, (4.6) is true, the proof is complete. □

From above fact, we can see that any optimal control $u(.)$ corresponding to x_0 must have the following form

$$u(t) = \begin{pmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_m(t) \end{pmatrix} \qquad \text{and} \qquad u_i(t) = \text{sign}\langle \theta, e^{-tA} b_i \rangle$$

where $\theta = -\zeta$ and $\zeta \in N_{\mathcal{R}(T)}$, and $B = [b_1, b_2, \ldots, b_m]$ is the m columns of B. Also observe that since we have defined any control which steer x_0 to the origin in time T is optimal, so we conclude

Corollary 4.8: If $x_0 \in \partial\mathcal{R}(T)$, then the control which is steer x_0 to the origin in time T is unique. Furthermore, it's is bang–bang and has the following form

$$u(t) = \begin{pmatrix} u_1(t) \\ u_2(t) \\ \vdots \\ u_m(t) \end{pmatrix} \qquad \text{and} \qquad u_i(t) = -\text{sign}\langle \zeta, e^{-tA} b_i \rangle$$

where $0 \neq \zeta$ is any normal vector in $N_{\mathcal{R}(T)}(x_0)$. And $B = [b_1, \ldots, b_m]$ are m-columns.

Surprisingly, the converse of this fact is also true and have many interesting consequences.

Theorem 4.9: If $x \in \mathcal{R}$ is steered to the origin in time T, and the corresponding control u^* has the form (where $B = [b_1, \ldots, b_m]$ are m-columns)

$$u^*(t) = \begin{pmatrix} u_1^*(t) \\ u_2^*(t) \\ \vdots \\ u_m^*(t) \end{pmatrix} \qquad \text{and} \qquad u_i^*(t) = -\text{sign}\langle \zeta, e^{-tA} b_i \rangle$$

where ζ is a non-zero vector in $H(A, B)$. Then we must have $x \in \partial\mathcal{R}(T)$ and $\zeta \in N_{\mathcal{R}(T)}(x)$, thus the above control u is optimal and unique.

Proof. We have $T(x) \leq T$, we need to prove $T(x) = T$, this implies $x \in \partial\mathcal{R}(T)$. Assume $T(x) = s < T$, then $x \in \partial\mathcal{R}(s)$, by using theorem 4.4. Now clearly x can be presented in form

$$x = -\int_0^T e^{-tA} B u^*(t)\, dt$$

For any $y \in \mathcal{R}(T)$, there exists a control $v \in \mathcal{U}_{ad}$ such that

$$y = -\int_0^T e^{-tA} B v(t)\, dt$$

Now we have

$$\begin{aligned}
\langle \zeta, y - x \rangle &= \left\langle \zeta, \int_0^T e^{-tA} B\Big[u^*(t) - v(t)\Big] \right\rangle \\
&= \sum_{i=1}^m \int_0^T \Big\langle \zeta, e^{-tA} b_i\big[u_i^*(t) - v_i(t)\big]\Big\rangle\, dt \\
&= \sum_{i=1}^m \int_0^T \Big(\big\langle \zeta, e^{-tA} b_i\big\rangle u_i^*(t) - \big\langle \zeta, e^{-tA} b_i\big\rangle v_i(t)\Big)\, dt \\
&= \sum_{i=1}^m \int_0^T \Big(-\big|\big\langle \zeta, e^{-tA} b_i\big\rangle\big| - \big\langle \zeta, e^{-tA} b_i\big\rangle v_i(t)\Big)\, dt \leq 0
\end{aligned}$$

Since $\mathcal{R}(s) \subset \mathcal{R}(T)$, we conclude

$$\langle \zeta, y - x \rangle \leq 0 \qquad \forall\, y \in \mathcal{R}(s)$$

thus, $\zeta \in N_{\mathcal{R}(s)}(x)$. Now using corollary 4.8, the unique control v^* which steer x to the origin in time s is also optimal and has the form

$$v^*(t) = \begin{pmatrix} v_1^*(t) \\ v_2^*(t) \\ \vdots \\ v_m^*(t) \end{pmatrix} \qquad \text{and} \qquad v_i^*(t) = -\text{sign}\langle \zeta, e^{-tA} b_i \rangle = u_i^*(t)$$

From this we easily see that $u^*(t) \equiv v^*(t)$ for all $t \in [0, s]$. So this implies that

$$\int_s^T e^{-tA} B u^*(t)\, dt = 0$$

Form this we have

$$0 = \int_s^T \left\langle \zeta, e^{-tA} B u^*(t) \right\rangle\, dt = \sum_{i=1}^m \int_s^T \left\langle \zeta, e^{-tA} b_i \right\rangle u_i^*(t)\, dt = -\sum_{i=1}^m \int_s^T \left| \left\langle \zeta, e^{-tA} b_i \right\rangle \right|\, dt$$

The final equation implies

$$\int_s^T \left| \left\langle \zeta, e^{-tA} b_i \right\rangle \right|\, dt = 0 \qquad \forall\, i = \overline{1, m}$$

i.e,

$$\left\langle \zeta, e^{-tA} b_i \right\rangle = 0 \qquad \forall\, i = \overline{1, m}, \forall\, t \in (s, T)$$

Doing similarly to the proof of theorem 3.4, it's give a contradiction since $\zeta \in H(A, B)$. Thus $T(x) = T$, i.e, $x \in \partial \mathbb{R}(T)$ and $\zeta \in N_{\mathcal{R}(T)}(x)$. □

In 4.5, we will use this result to prove a nice fact, that is every point is optimal if $\mathcal{R} = \mathbb{R}^n$.

4.3 Application to Rocket railroad car problem

4.3.1 The switching time of the bang–bang optimal control

Now we will extract the interesting fact that the maximum principle enable us to know the switching time of the bang–bang optimal control, let's recall the rocket rail road car problem

$$\begin{cases} z'(t) & = A \cdot z(t) + b \cdot u(t) \\ z(0) & = (x_0, v_0)^T \end{cases}$$

where $u(t) \in U = [-1,1]$ and

$$z(t) = \begin{pmatrix} x(t) \\ v(t) \end{pmatrix} \qquad A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Let's consider the controllability matrix

$$G = \left[\begin{pmatrix} 0 \\ 1 \end{pmatrix}; \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Note that because rank $G = 2$, and also the condition Re $\lambda \leq 0$ for eigenvalue λ of A is satisfied since $\lambda = \pm i$. Thus we have $\mathcal{R} = H(A, b) = \mathbb{R}^2$. Let u^* be a bang–bang optimal control for z_0, according to the maximum principle, then there exists $\theta \in \mathbb{R}^2$ such that

$$\theta^T e^{-sA} b u^*(s) = \max_{|\lambda| \leq 1} \{\theta^T e^{-sA} b \lambda\} \qquad \forall\, s \in [0, T(z_0)] \tag{4.10}$$

Let's compute e^{-sA}, observe that $A^0 = I_2$, $A^n = 0$ for all $n \geq 2$, so

$$e^{-sA} = I - sA = \begin{pmatrix} 1 & -s \\ 0 & 1 \end{pmatrix}$$

Assume $\theta = (\theta_1, \theta_2)$ then

$$\theta^T e^{-sA} b = (\theta_1, \theta_2) \begin{pmatrix} 1 & -s \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -s\theta_1 + \theta_2$$

So by (4.10) we have

$$(-s\theta_1 + \theta_2) u^*(s) = \max_{|\lambda| \leq 1} \{(-s\theta_1 + \theta_2)\lambda\} \Longrightarrow u^*(s) = \operatorname{sign}(-s\theta_1 + \theta_2)$$

where sign is the sign function, i.e

$$\operatorname{sign}(x) = \begin{cases} 1 & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}$$

From this fact, we easily see that the bang–bang optimal control u^* switches at most once. Further more, if $\theta_1 = 0$ then u^* is a constant, 1 or -1.

Now from this, the geometric solution of rocket problem that we had introduced before in 2.3.3 must have been a bang–bang optimal control.

4.3.2 The reachable sets

Now we want to draw the reachable set $\mathcal{R}(t)$ of the rocket rail road car problem.

Note that $\mathcal{R}(t)$ is compact, strictly convex in $\mathbb{R}^2$. So to know $\mathcal{R}^2$, we only need to know the boundary $\partial \mathcal{R}(t)$.

Let's find the boundary $\partial\mathcal{R}(t)$. A point $z_0 = (x_0, v_0) \in \partial\mathcal{R}(t)$ if and only if

$$z_0 = -\int_0^t e^{-sA} b u(s)\, ds$$

where $u(s)$ is a bang–bang optimal control, i.e u switches at most once. So such an bang–bang optimal control like this must have one of two forms

$$(a) \quad u_a(s) = \begin{cases} 1 & \text{on } [0, t_0] \\ -1 & \text{on } (t_0, t] \end{cases} \qquad \text{or} \qquad (b) \quad u_b(s) = \begin{cases} -1 & \text{on } [0, t_0] \\ 1 & \text{on } (t_0, t] \end{cases}$$

So the set of all point on $\partial\mathcal{R}(t)$ is compatible to the set of all admissible control u of above forms. Let's finding z_0 in each case

$$z_0 = -\int_0^t e^{-sA} b u(s) ds = -\int_0^t \begin{pmatrix} 1 & -s \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(s)\, ds = -\int_0^t \begin{pmatrix} -s \\ 1 \end{pmatrix} u(s)\, ds$$

(a) If $u = u_a$ for $t_0 \in [0, t]$, we have

$$\begin{aligned} x_0 &= -\int_0^t -s u(s)\, ds = -\int_0^{t_0} -s\, ds - \int_{t_0}^t s\, ds = t_0^2 - \frac{t^2}{2} \\ v_0 &= -\int_0^t u(s)\, ds = -\int_0^{t_0} 1\, ds - \int_{t_0}^t -1\, ds = t - 2t_0 \end{aligned}$$

So (x_0, v_0) must lie on the curve ($t_0 \in [0, t]$)

$$(C_a): \qquad x_0 + \frac{t^2}{2} = \left(\frac{t - v_0}{2}\right)^2 \qquad \text{i.e} \qquad (C_a): \qquad v_0^2 = 2tv_0 + 4x_0 + t^2$$

(b) If $u = u_b$ for $t_0 \in [0, t]$, we have

$$\begin{aligned} x_0 &= -\int_0^t -s u(s)\, ds = -\int_0^{t_0} s\, ds - \int_{t_0}^t -s\, ds = -t_0^2 + \frac{t^2}{2} \\ v_0 &= -\int_0^t u(s)\, ds = -\int_0^{t_0} -1\, ds - \int_{t_0}^t 1\, ds = -t + 2t_0 \end{aligned}$$

So (x_0, v_0) must lie on the curve ($t_0 \in [0, t]$)

$$(C_b): \qquad -x_0 + \frac{t^2}{2} = \left(\frac{t + v_0}{2}\right)^2 \qquad \text{i.e} \qquad (C_b): \qquad v_0^2 = -2tv_0 - 4x_0 + t^2$$

It's easy to see that (C_a) and (C_b) is symmetric through the origin.

For example, let's consider the geometric presentation of $\partial\mathcal{R}(2)$

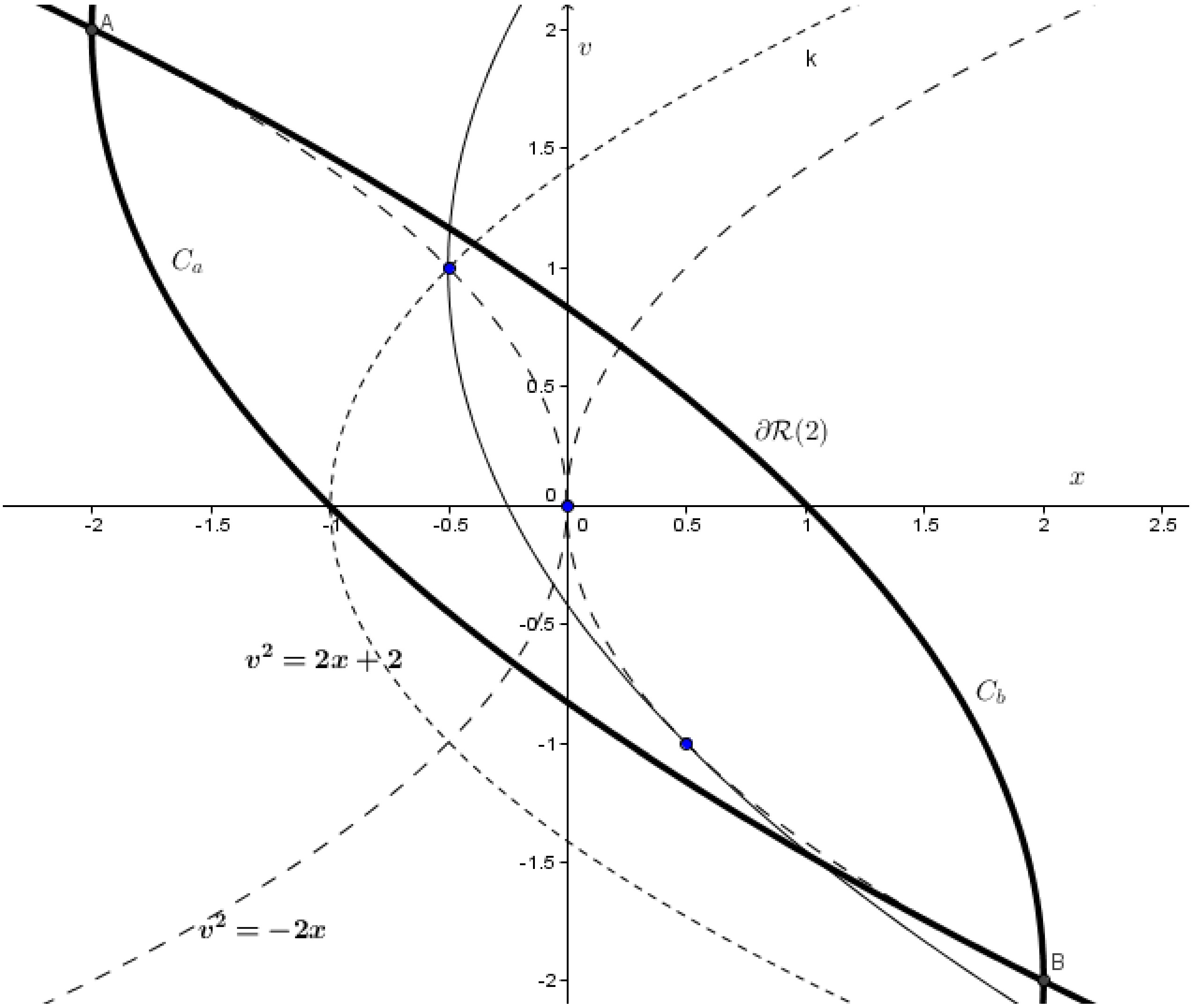


Figure 5: The boundary of $\mathcal{R}(2)$

In this cases, the point $A = (x_0, v_0) = (-2, 2)$ which is respect to the control $u = -1$, and the trajectory which is steer A to the origin just is the parabola $v^2 = -2x$. Another observation, consider $(-1, 0)$ which is respect to the control

$$u(s) = \begin{cases} 1 & \text{on } [0,1] \\ -1 & \text{on } (1,2] \end{cases}$$

Then the trajectory which steer $(-1, 0)$ to the origin is first go on the parabola $v^2 = 2x + 2$, then when first meet the parabola $v^2 = -2x$, switching to it parabola and go to the origin.

We will see some geometric presentation of $\mathcal{R}(t)$ when t is increasing.

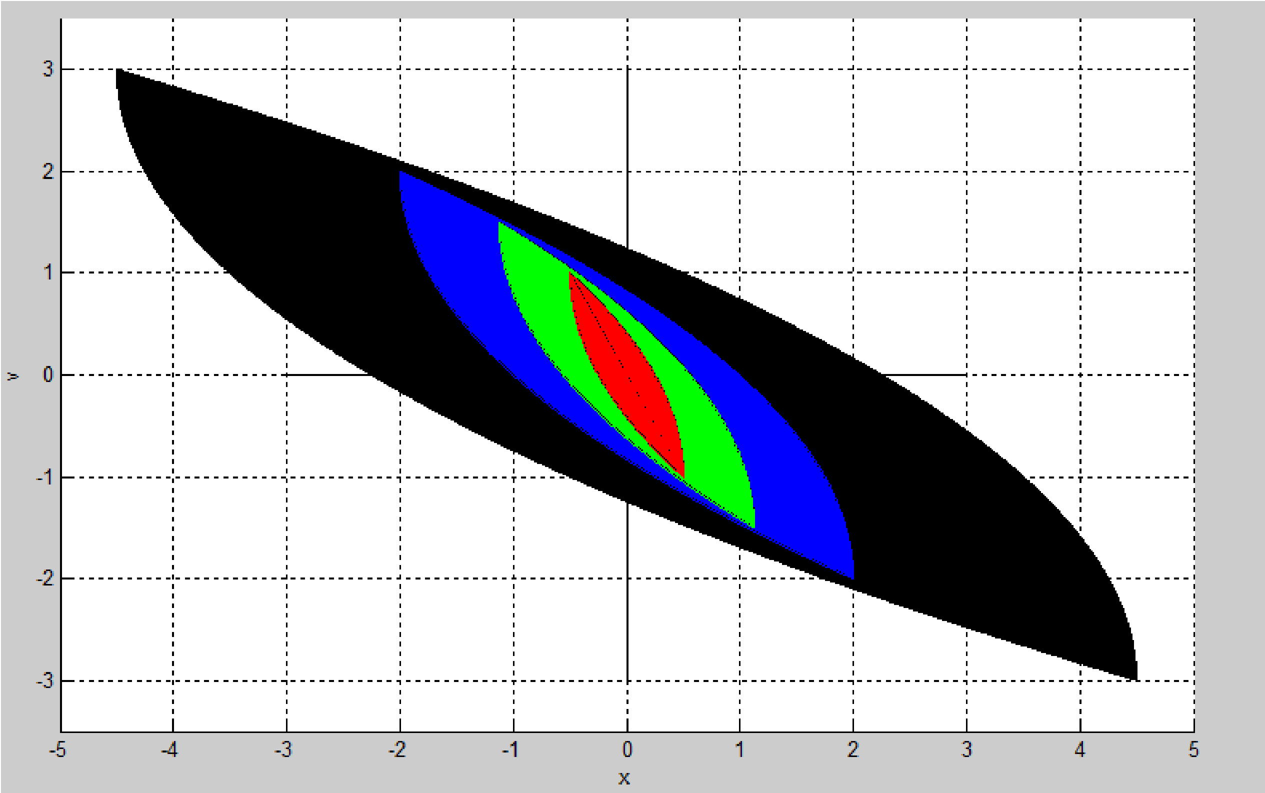


Figure 6: Reachable set for $t = 3, 2, 1.5, 1$

We can see that for $t \longrightarrow 0$, the reachable set $\mathcal{R}(t)$ is decreasing but always contain 0 as an interior point, as we had proved before.

4.3.3 The minimum time function

Now we can find the formula for the minimum time function for the rocker rail road car problem, by using above results.

Let $z_0 = (x_0, v_0) \in \mathbb{R}^2$, the the minimum time $T(z_0)$ is exactly the time t, where (x_0, v_0) lie on the boundary $\partial\mathcal{R}(t)$, in other words, we can see that (x_0, v_0) must lie on either $(C_a^t) : t^2 + 2tv_0 + 4x_0 - v_0^2 = 0$ or $(C_b^t) : t^2 - 2tv_0 - 4x_0 - v_0^2 = 0$.

So given (x_0, v_0), we only need to solve two equations

$$(C_a): \qquad t^2 + 2tv_0 + 4x_0 - v_0^2 = 0 \qquad \Delta'_a = 2(v_0^2 - 2x_0) \tag{4.11}$$

$$(C_b): \qquad t^2 - 2tv_0 - 4x_0 - v_0^2 = 0 \qquad \Delta'_b = 2(v_0^2 + 2x_0) \tag{4.12}$$

Then choosing the minimum positive solution of these solutions, it's must be $T(x_0, v_0)$. We also note that at least one of these equations must have positive solution, since

$$\Delta'_a + \Delta'_b = 4v_0^2 \geq 0$$

so at least one equation must have solution, and the existence of positive solution following

- If $\Delta'_a \geq 0$. If $\Delta'_a > 0$, then clearly $4x_0 - v_0^2 < 0$, so the equation (4.11) has two distinct solutions, one of them is positive clearly, given by

$$t = -v_0 + \sqrt{2(v_0^2 - 2x_0)}$$

  If $\Delta'_a = 0$, i.e $v_0^2 = 2x_0$, so $x_0 \geq 0$, then (4.11) become $t^2 + 2tv_0 + v_0^2 = (t + v_0)^2 = 0$

  - If $v_0 < 0$, then $t = -v_0 > 0$ is a positive solution.
  - If $v_0 > 0$, then (4.12) become $t^2 - 2tv_0 - 3v_0^2 = 0$, i.e $(t + v_0)(v - 3v_0) = 0$, so $t = 3v_0 > 0$ is a positive solution. It's can be also given by

$$t = v_0 + \sqrt{2(v_0^2 + 2x_0)} = 3v_0$$

  - If $v_0 = 0$, then $x_0 = 0$ and we have nothing to do in case $(x_0, v_0) \equiv (0, 0)$.

- If $\Delta'_b \geq 0$. If $\Delta'_b > 0$, then clearly $-4x_0 - v_0^2 < 0$, so the equation (4.12) has two distinct solutions, one of them is positive clearly, given by

$$t = v_0 + \sqrt{2(v_0^2 + 2x_0)}$$

  If $\Delta'_b = 0$, i.e $v_0^2 = -2x_0$, so $x_0 \leq 0$ then (4.12) become $t^2 - 2tv_0 + v_0^2 = (t - v_0)^2 = 0$

  - If $v_0 > 0$, then $t = v_0 > 0$ is a positive solution.
  - If $v_0 < 0$, then (4.11) become $t^2 + 2tv_0 - 3v_0^2 = 0$, i.e $(t - v_0)(v + 3v_0) = 0$, so $t = -3v_0 > 0$ is a positive solution. It's can be also given by

$$t = -v_0 + \sqrt{2(v_0^2 - 2x_0)} = -3v_0$$

  - If $v_0 = 0$, then $x_0 = 0$ and we have nothing to do in case $(x_0, v_0) \equiv (0, 0)$.

To summarize this fact, given (x_0, v_0), we have

$$t = T(x_0, v_0) = \begin{cases} -v_0 + \sqrt{2(v_0^2 - 2x_0)} & \text{if} \quad 2x_0 < -v_0|v_0| \\ v_0 + \sqrt{2(v_0^2 + 2x_0)} & \text{if} \quad 2x_0 \geq -v_0|v_0| \end{cases}$$

Now we have a geometric presentation of this formula.

From this presentation, we can see that the minimum time function is continuous as we had proved before, but there some point at that it's not differentiable.

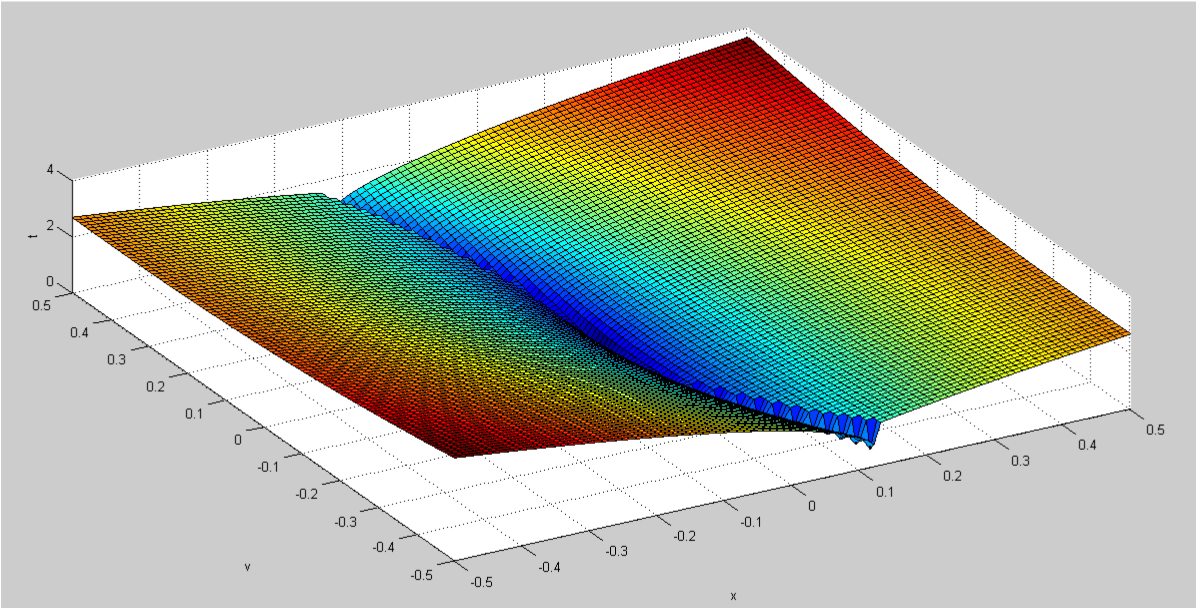


Figure 7: Geometric presentation of the minimum time function

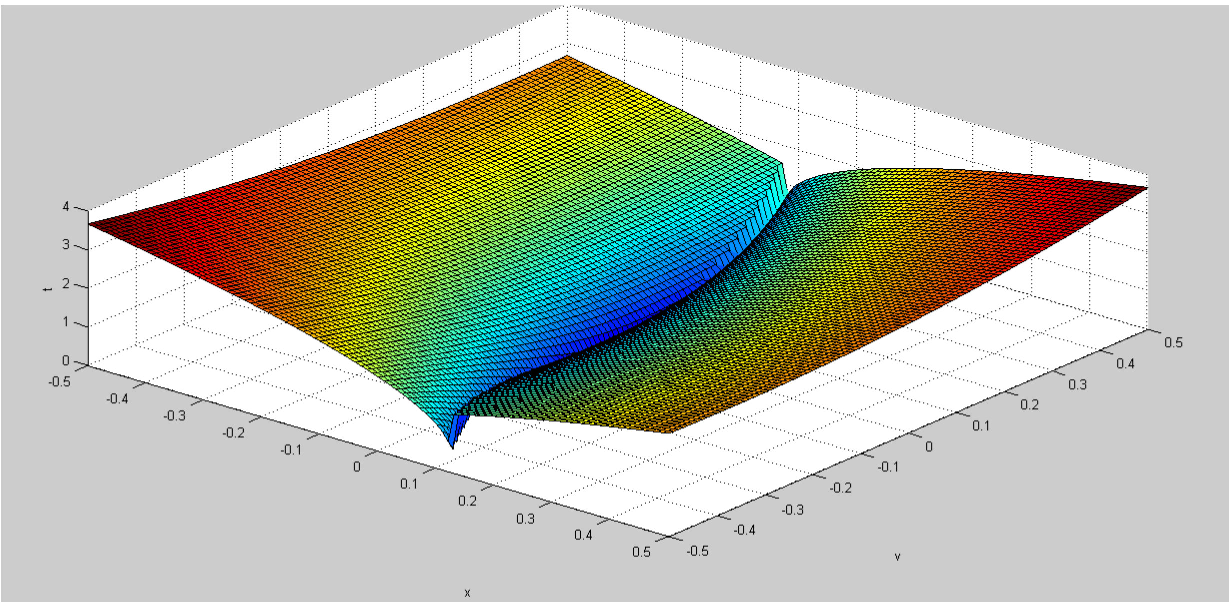


Figure 8: Geometric presentation of the minimum time function

We can see that along to the curves

$$\begin{aligned} v^2 &= -2x, \qquad v \geq 0 \\ v^2 &= \ \ 2x \qquad\ \ v \leq 0 \end{aligned}$$

the minimum time function of the rocket car problem is not smooth. So this function is not smooth in whole domain $\mathcal{R}$, and even non-Lipschitz, as we will prove later. Furthermore, in 5.2.1 we will prove that the union of these curves contains all points at which T is non–Lipschitz.

4.4 Application to the Harmonic oscillator

4.4.1 The switching time of the bang–bang optimal control

Let's consider the controllability matrix

$$G = \left[\begin{pmatrix} 0 \\ 1 \end{pmatrix} ; \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right] = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Since rank $G = 2$, and also eigenvalues of A are $\pm i$, so we have $\mathcal{R} = H(A, b) = \mathbb{R}^2$. Let u^* be a bang–bang optimal control for z_0, according to the maximum principle, there exists $\theta \in \mathbb{R}^2$ such that

$$\theta^T e^{-sA} b u^*(s) = \max_{|\lambda| \leq 1} \{\theta^T e^{-sA} b \lambda\} \qquad \forall\, s \in [0, T(z_0)] \tag{4.13}$$

Let's compute e^{-sA}, we have

$$A^0 = I_2 \qquad A^1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad A^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -I_2 \qquad A^3 = -A$$

So by induction, it's easy to see that $A^{4k+r} = A^r$ for all $k \in \mathbb{N}$ and $r = 0, 1, 2, 3$. So we have

$$\begin{aligned} e^{sA} &= \sum_{n=0}^{\infty} \frac{s^n}{n!} A^n = \sum_{k=0}^{\infty} \frac{s^{2k}}{(2k)!} A^{2k} + \sum_{k=0}^{\infty} \frac{s^{2k+1}}{(2k+1)!} A^{2k+1} = I_2 \left(\sum_{k=0}^{\infty} \frac{(-1)^k s^{2k}}{(2k)!} \right) + A \left(\sum_{k=0}^{\infty} \frac{(-1)^k s^{2k+1}}{(2k+1)!} \right) \\ &= (\cos s)\, I_2 + (\sin s)\, A = \begin{pmatrix} \cos s & \sin s \\ -\sin s & \cos s \end{pmatrix} \end{aligned}$$

So we have $e^{-sA} = \begin{pmatrix} \cos(-s) & \sin(-s) \\ -\sin(-s) & \cos(-s) \end{pmatrix} = \begin{pmatrix} \cos s & -\sin s \\ \sin s & \cos s \end{pmatrix}$, assume $\theta = (\theta_1, \theta_2)^T$ then

$$\theta^T e^{-sA} b = (\theta_1, \theta_2) \begin{pmatrix} \cos s & -\sin s \\ \sin s & \cos s \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\theta_1 \sin s + \theta_2 \cos s$$

By (4.13) we have

$$\theta^T e^{-sA} b u^*(s) = \max_{|\lambda| \leq 1} \{\theta^T e^{-sA} b \lambda\} \Longrightarrow u^*(s) = \operatorname{sign}\left(-\theta_1 \sin s + \theta_2 \cos s\right)$$

Now, since the vector θ is choosing to be unit, i.e $\|\theta\|^2 = \theta_1^2 + \theta_2^2 = 1$, so for simplify further we may choose δ such that $\cos\delta = -\theta_1$ and $\sin\delta = \theta_2$. Then

$$u^*(s) = \operatorname{sign}\left(-\theta_1 \sin s + \theta_2 \cos s\right) = \operatorname{sign}\left(\cos\delta \sin s + \sin\delta \cos s\right) = \operatorname{sign}\left(\sin(s + \delta)\right)$$

From this, we ca deduce that when time s is increase, then every s pass through π–units of time, then u^* will be switch between +1 and -1.

4.4.2 Geometric interpretation

We will figure out the geometric solution by this motion, by considering the case $u \equiv 1$ and $u \equiv -1$.

Case $u^* \equiv 1$. Our equation (2.24) becomes

$$\begin{cases} x'(t) &= v(t) \\ v'(t) &= -x(t)+1 \end{cases} \Longrightarrow \frac{\partial}{\partial t}\Big[(-x(t)+1)^2 + v^2(t)\Big] = 0$$

So the motion must satisfy $(x(t)-1)^2 + v^2(t) = r^2$ for some constant r. I.e the trajectory must lie on a circle center at (1,0).

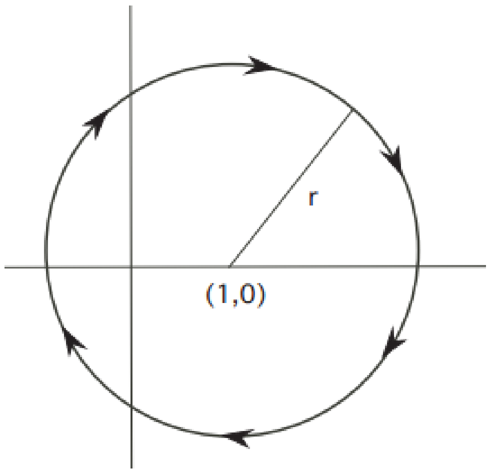


Figure 9: The trajectory in case $u^* \equiv 1$

Indeed, if $u \equiv 1$ in $[0, t_0]$, then the trajectory in $[0, t_0]$ has the form

$$z(t) = e^t z_0 + \int_0^t e^{(t-s)A} B u^*(s)\, ds \Longrightarrow \begin{pmatrix} x(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} x_0 \cos t + v_0 \sin t + 1 - \cos t \\ -x_0 \sin t + v_0 \cos t + \sin t \end{pmatrix}$$

for all $t \in [0, t_0]$. This fact explains why is the orientation of this trajectory like this picture.

Case $u^* \equiv -1$. Our equation (2.24) becomes

$$\begin{cases} x'(t) &= v(t) \\ v'(t) &= -x(t)-1 \end{cases} \Longrightarrow \frac{\partial}{\partial t}\Big[(x(t)+1)^2 + v^2(t)\Big] = 0$$

So the motion must satisfy $(x(t)+1)^2 + v^2(t) = r'^2$ for some constant r'. I.e the trajectory must lie on a circle center at (-1,0).

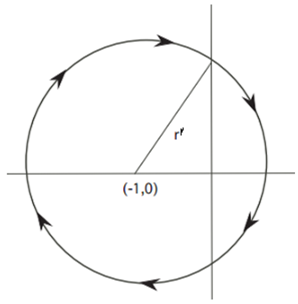


Figure 10: The trajectory in case $u^* \equiv -1$

Indeed, if $u \equiv -1$ in $[0, t_0]$, then the trajectory in $[0, t_0]$ has the form

$$z(t) = e^{tA}z_0 + \int_0^t e^{(t-s)A}Bu^*(s)\, ds \Longrightarrow \begin{pmatrix} x(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} x_0 \cos t + v_0 \sin t - 1 + \cos t \\ -x_0 \sin t + v_0 \cos t - \sin t \end{pmatrix}$$

for all $t \in [0, t_0]$. This fact explains why is the orientation of this trajectory like this picture.

Summary. In summary, to get to the origin we must switch our control u^* back and forth between the values ± 1, causing the trajectory to switch between lying on circles centered at $(\pm 1, 0)$. The switches occur each π units of time. Too understand deeply, we construct an control like this when $z_0 = (\cos 4 + 3, \sin 4)^T$. First, assume control $u = 1$, the trajectory lies on the circle center as $(1, 0)$ pass through z_0, with the orientation like above figure. Until the time $t = \pi$, the control switches to 1 and the trajectory switches to the circle center as $(-1, 0)$ pass through the point $z(\pi)$, as following picture

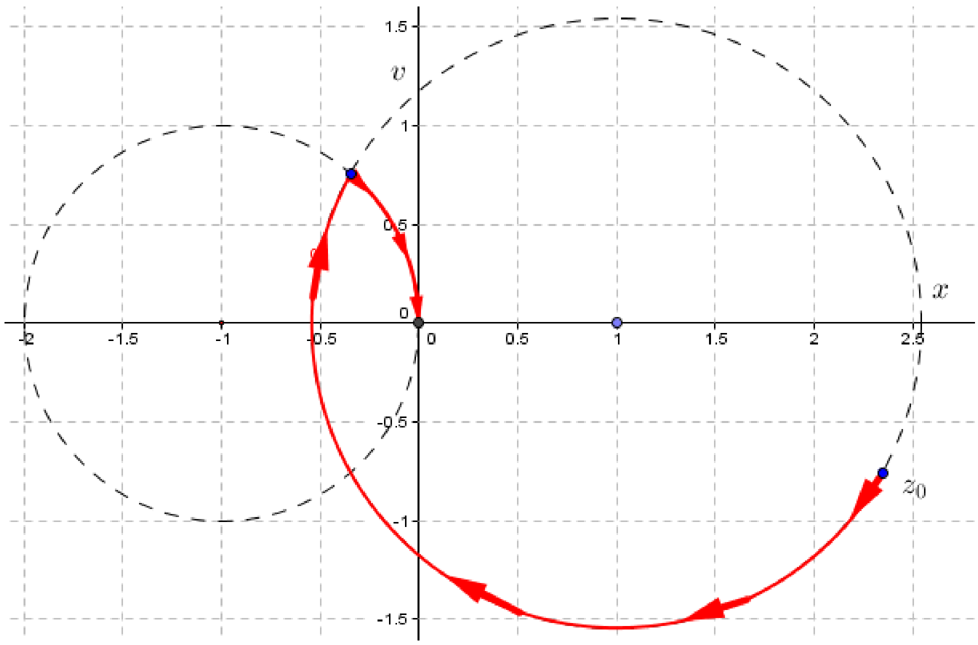


Figure 11: The trajectory steers $(\cos 4 + 3, \sin 4)$ to the origin in time 4s

Another example by $z_0 = \begin{pmatrix} -2\cos 1 \\ -2\sin 1 \end{pmatrix}$ in time $t = \pi$. We have

$$\begin{pmatrix} -2\cos 1 \\ -2\sin 1 \end{pmatrix} \xrightarrow[1s]{u=1} \begin{pmatrix} -1-\cos 1 \\ \sin 1 \end{pmatrix} \xrightarrow[(\pi-1)s]{u=-1} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

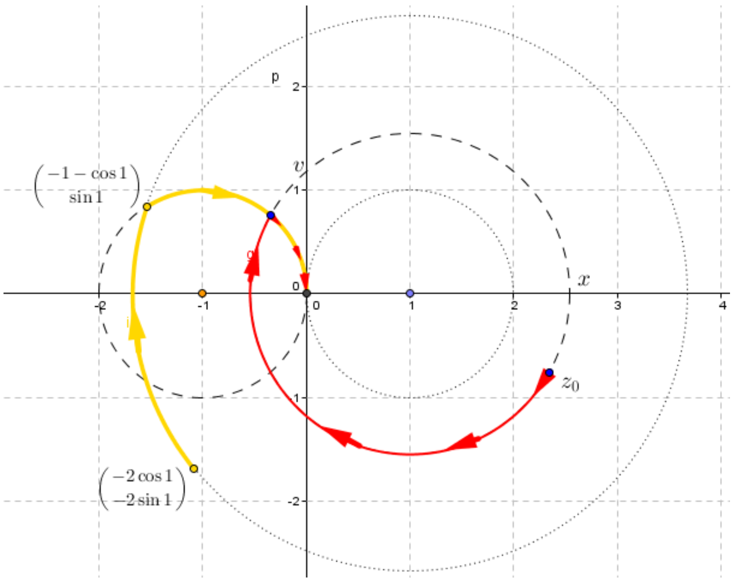


Figure 12: The trajectory steers $(-2\cos 1, -2\sin 1)$ to the origin in time πs

4.4.3 The reachable sets

Now we want to draw the reachable set $\mathcal{R}(t)$ of this motion, note that since $\mathcal{R} = \mathbb{R}^2$, so there nothing to draw about $\mathcal{R}$. Note that $\mathcal{R}(t)$ is compact, strictly convex in $\mathbb{R}^2$. So to know $\mathcal{R}^2$, we only need to know the boundary $\partial\mathcal{R}(t)$.
Let's find the boundary $\partial\mathcal{R}(t)$. A point $z_0 = (x_0, v_0) \in \partial\mathcal{R}(t)$ if and only if

$$z_0 = -\int_0^t e^{-sA} b u(s)\, ds$$

where $u(s)$ is a bang–bang optimal control. From the fact that $u^*(s) = \text{sign}\,(\sin(s+\delta))$, given a time $t > 0$, the geometric property of $\mathcal{R}(t)$ depend on the relation about t and π.

The reachble set $\mathcal{R}(\pi)$ First we draw $\mathcal{R}(t)$ for $t = \pi$. Consider vertexs, it's consist of two kinds of points, one of them is respect to the control $u = 1$ in $[0, \pi]$, another is respect to the control $u = -1$ in $[0, \pi]$. And there are two cases about a vertexs $z_0 \in \partial\mathcal{R}(\pi)$ with no switch time during the motion

$$z_0 = -\int_0^\pi e^{-sA} b\, ds = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \qquad \text{or} \qquad z_0 = \int_0^\pi e^{-sA} b\, ds = \begin{pmatrix} -2 \\ 0 \end{pmatrix}$$

Now, all points of $\mathcal{R}(\pi)$ except two vertexs has exactly one switch times. I.e it must has once of two form

$$(a) \quad u_a(s) = \begin{cases} 1 & \text{on } [0,t] \\ -1 & \text{on } (t,\pi] \end{cases} \qquad \text{or} \qquad (b) \quad u_b(s) = \begin{cases} -1 & \text{on } [0,t] \\ 1 & \text{on } (t,\pi] \end{cases}$$

So the set of all point on $\partial\mathcal{R}(\pi)$ is compatible to the set of all admissible control u of above forms. Let's finding z_0 in each case

$$z_0 = -\int_0^t e^{-sA} b u(s) ds = -\int_0^\pi \begin{pmatrix} \cos s & -\sin s \\ \sin s & \cos s \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(s)\, ds = \int_0^\pi \begin{pmatrix} \sin s \\ -\cos s \end{pmatrix} u(s)\, ds \tag{4.14}$$

(a) $u = 1$ on $[0,t]$ and $u = -1$ on $[t,\pi]$, we have

$$z_0^a = \begin{pmatrix} -\cos s|_0^t + \cos s|_t^\pi \\ -\sin s|_0^t + \sin s|_t^\pi \end{pmatrix} = \begin{pmatrix} -2\cos t \\ -2\sin t \end{pmatrix}$$

So we have $z_0^a = (x_0, v_0)$ must live on the curve $(C_a) : v = -\sqrt{4-x^2}$, $v \le 0$.

(b) $u = -1$ on $[0,t]$ and $u = 1$ on $[t,\pi]$, we have

$$z_0^b = \begin{pmatrix} \cos s|_0^t - \cos s|_t^\pi \\ \sin s|_0^t - \sin s|_t^\pi \end{pmatrix} = \begin{pmatrix} 2\cos t \\ 2\sin t \end{pmatrix}$$

So we have $z_0^b = (x_0, v_0)$ must live on the curve $(C_b) : v = \sqrt{4-x^2}$, $v \ge 0$.

It's easy to see that they are symmetric through the origin as we have prove before. Finally, we draw the boundary of $\mathcal{R}(\pi)$ as the circle center at $(0,0)$ has the radius of 2.

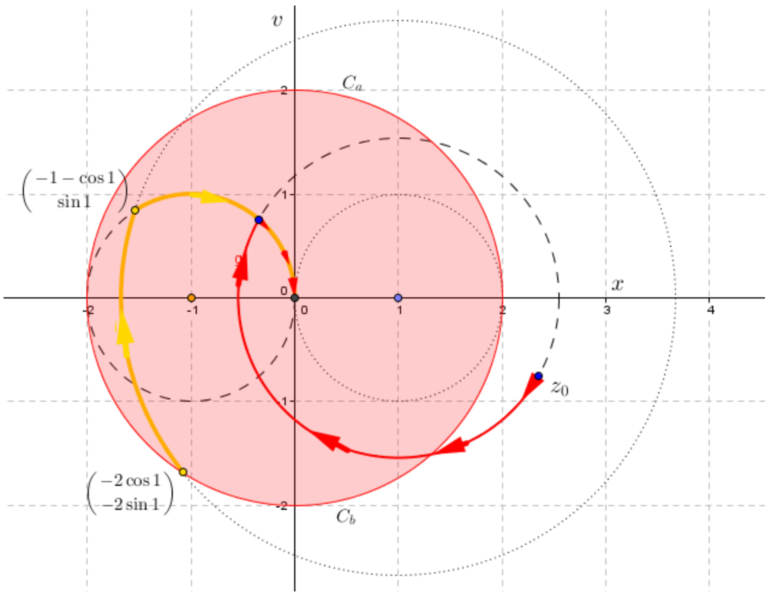


Figure 13: The reachable set $\mathcal{R}(\pi)$ with boundary consist of two semicircles C_a and C_b

The reachble set $\mathcal{R}(t)$ where $t < \pi$ Doing similarly from above steps, since in $[0,t]$ where $t<\pi$, the maximum principle says that the motion switches at most once.

Now, all points of $\mathcal{R}(t)$ has exactly one switch times. I.e it must has once of two form

$$(a)\quad u_a(s) = \begin{cases} 1 & \text{on } [0,t_0] \\ -1 & \text{on } (t_0,t] \end{cases} \qquad \text{or} \qquad (b)\quad u_b(s) = \begin{cases} -1 & \text{on } [0,t_0] \\ 1 & \text{on } (t_0,t] \end{cases}$$

So the set of all point on $\partial\mathcal{R}(\pi)$ is compatible to the set of all admissible control u of above

forms. Let's finding z_0 in each case

$$z_0 = -\int_0^t e^{-sA} bu(s) ds = -\int_0^t \begin{pmatrix} \cos s & -\sin s \\ \sin s & \cos s \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(s)\, ds = \int_0^t \begin{pmatrix} \sin s \\ -\cos s \end{pmatrix} u(s)\, ds$$

(a) $u=1$ on $[0,t_0]$ and $u=-1$ on $[t_0,t]$, we have

$$z_0^a = \begin{pmatrix} -\cos s|_0^{t_0} + \cos s|_{t_0}^t \\ -\sin s|_0^{t_0} + \sin s|_{t_0}^t \end{pmatrix} = \begin{pmatrix} -2\cos t_0 + 1 + \cos t \\ -2\sin t_0 + \sin t \end{pmatrix}$$

So we have $z_0^a = (x_0, v_0)$ must live on the curve

$$(C_a) : \left(x_0 - (1+\cos t)\right)^2 + (v_0 - \sin t)^2 = 4$$

where

$$v_0 - \sin t = -2\sin t \le 0$$

since $t \in [0,\pi)$.

(b) $u=-1$ on $[0,t_0]$ and $u=1$ on $[t_0,t]$, we have

$$z_0^b = \begin{pmatrix} \cos s|_0^{t_0} - \cos s|_{t_0}^t \\ \sin s|_0^{t_0} - \sin s|_{t_0}^t \end{pmatrix} = \begin{pmatrix} 2\cos t_0 - 1 - \cos t \\ 2\sin t_0 - \sin t \end{pmatrix}$$

So we have $z_0^a = (x_0, v_0)$ must live on the curve

$$(C_b) : \left(x_0 - (-1-\cos t)\right)^2 + \left(v_0 - (-\sin t)\right)^2 = 4$$

where

$$v_0 + \sin t = 2\sin t \le 0$$

since $t \in [0,\pi)$.

It's easy to see that they are also symmetric through the origin as we have prove before. Finally, we draw the boundary of $\mathcal{R}(t)$ as following, two vertexs are respect to $t_0 = 0$ and $t_0 = t$.

$$\begin{pmatrix} -1+\cos t \\ \sin t \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 1-\cos t \\ -\sin t \end{pmatrix}$$

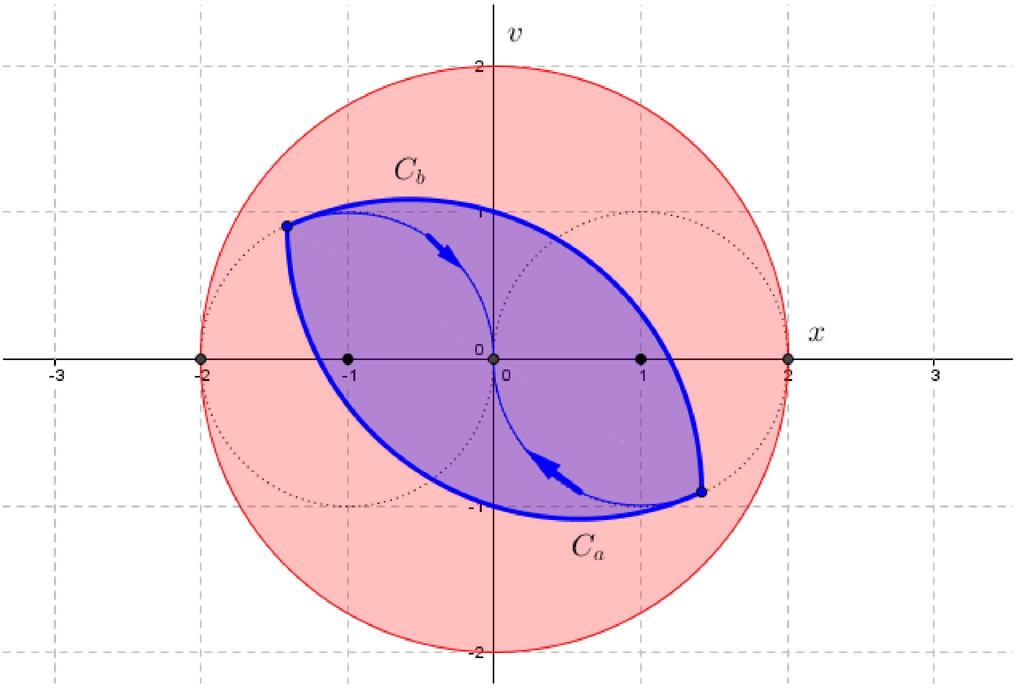


Figure 14: The reachable set $\mathcal{R}(2)$ $(2 < \pi)$

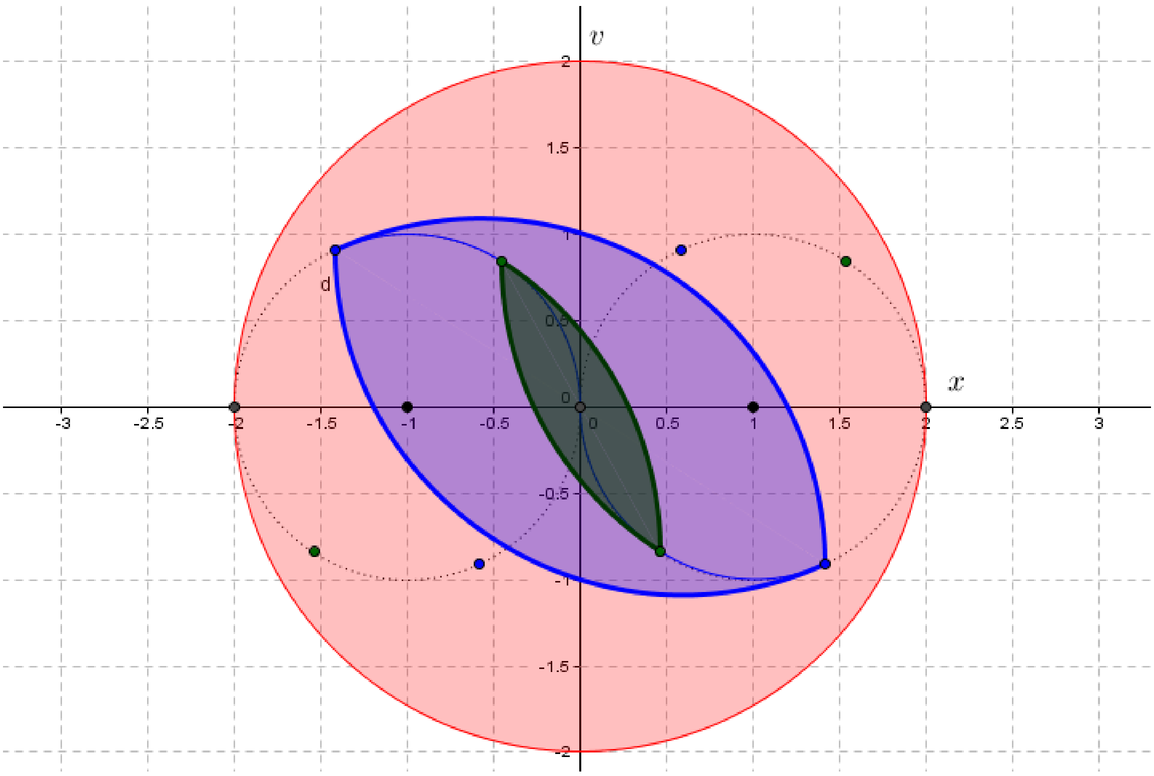


Figure 15: The reachable set $\mathcal{R}(t)$ where $t = 1, 2, \pi$

The reachble set $\mathcal{R}(t)$ where $\pi < t \le 2\pi$ When $\pi < t \le 2\pi$, then also from maximum principle, we conclude that on the boundary, the motion switches at most two times. We have four cases

- Case 1. $0 \le t_0 \le t - \pi$ and

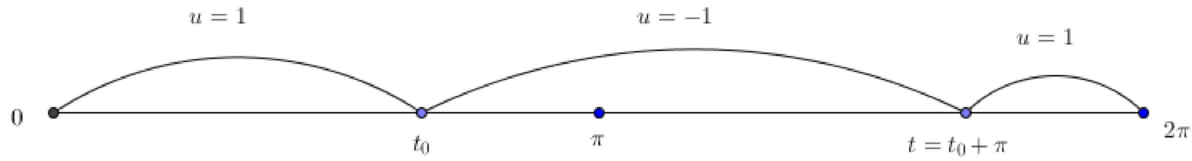


Figure 16: The control u in case 1.

Easily from (4.14) we have

$$(C1) \qquad z_0 = \begin{pmatrix} x_0 \\ v_0 \end{pmatrix} = \begin{pmatrix} -4\cos t_0 + 1 - \cos t \\ -4\sin t_0 - \sin t \end{pmatrix} \qquad t_0 \in [0, t - \pi]$$

This is an arc center as $(1 - \cos t, -\sin t)$ with radius 4.

- Case 2. $0 \le t_0 \le t - \pi$ and

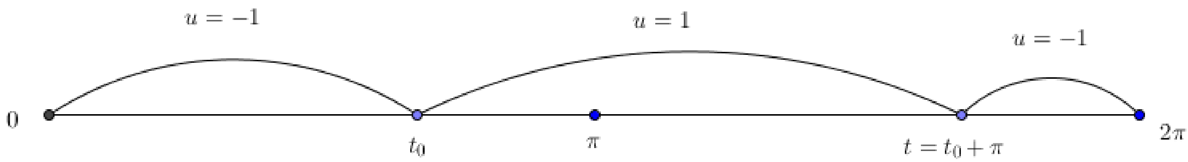


Figure 17: The control u in case 2.

Easily from (4.14) we have

$$(C2) \qquad z_0 = \begin{pmatrix} x_0 \\ v_0 \end{pmatrix} = \begin{pmatrix} 4\cos t_0 - 1 + \cos t \\ 4\sin t_0 + \sin t \end{pmatrix} \qquad t_0 \in [0, t - \pi]$$

This is an arc center as $(-1 + \cos t, \sin t)$ with radius 4.

- Case 3. $t - \pi \le t_0 \le \pi$ and

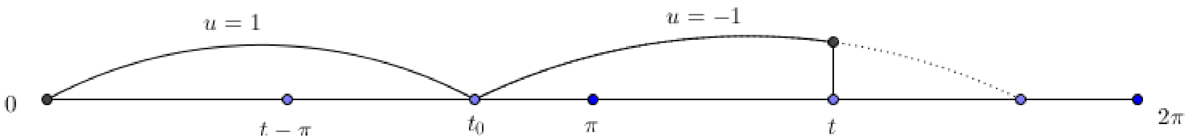


Figure 18: The control u in case 3.

Easily from (4.14) we have

$$(C3) \qquad z_0 = \begin{pmatrix} x_0 \\ v_0 \end{pmatrix} = \begin{pmatrix} -2\cos t_0 + \cos t + 1 \\ -2\sin t_0 + \sin t \end{pmatrix} \qquad t_0 \in [t - \pi, \pi]$$

This is an arc center as $(1 + \cos t, \sin t)$ with radius 2.

- Case 4. $t - \pi \le t_0 \le \pi$ and

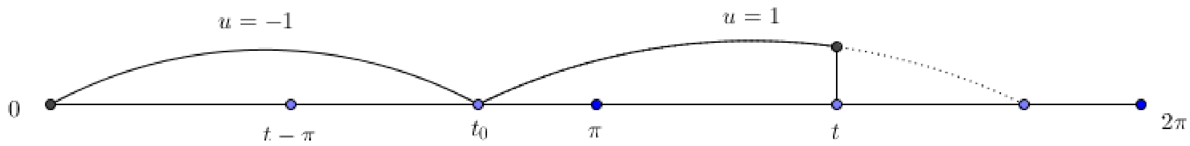


Figure 19: The control u in case 4.

Easily from (4.14) we have

$$(C4) \qquad z_0 = \begin{pmatrix} x_0 \\ v_0 \end{pmatrix} = \begin{pmatrix} 2\cos t_0 - \cos t - 1 \\ 2\sin t_0 - \sin t \end{pmatrix} \qquad t_0 \in [t-\pi, \pi]$$

This is an arc center as $(-1-\cos t, -\sin t)$ with radius 2.

It's easy to see that $(C1), (C2)$ are symmetric through the origin, and the same to $(C3)$ and $(C4)$. They make a closed curve, and $\mathcal{R}(t)$ is follow

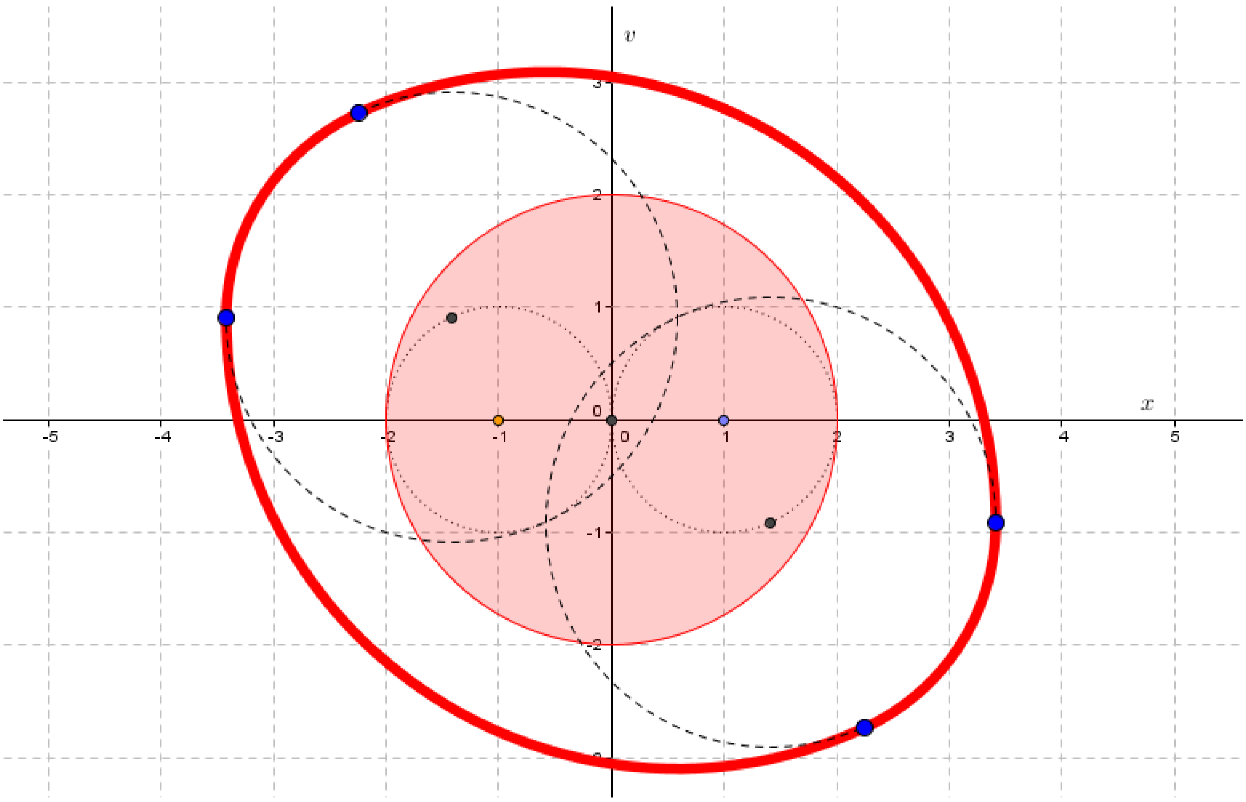


Figure 20: The boundary of the reachable set $\mathcal{R}(\pi+2)$.

As we can seen, the boundary of $\mathcal{R}(t)$ seem to be "smooth" when $t \geq \pi$, in the following picture

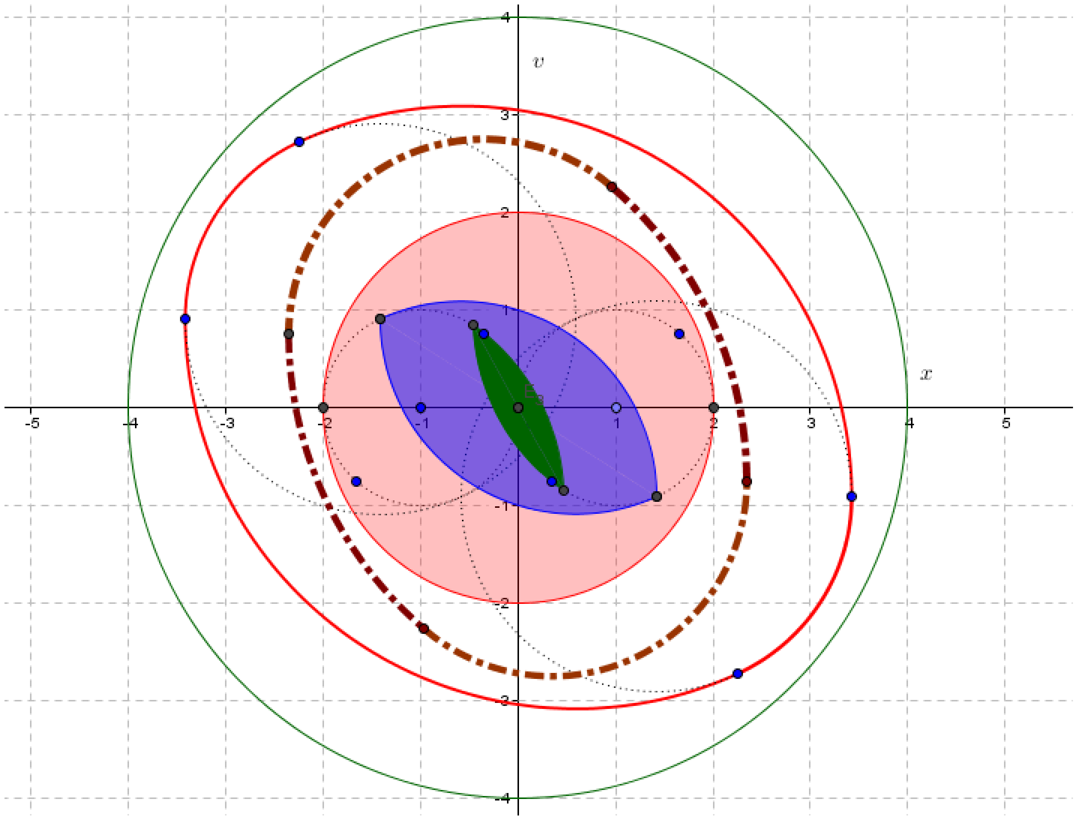


Figure 21: The boundary of the reachable set $\mathcal{R}(t)$ where $t = 1, 2, \pi, 4, 2 + pi, 2\pi$.

The reachable set $\mathcal{R}(t)$ **where** $2\pi < t < \infty$ As we have seen from two above cases, we just doing similarly to obtain that there exists a unique $n \in \mathbb{N}$ such that $t \in (n\pi, (n+1)\pi]$, and there are almost four types of optimal control u^*. This make into two pairs, in which pair two optimal controls are symmetric through 0.

Type 1. u has $n + 1$ switching times. This type consist of two cases, which are symmetric through 0, with $0 < t_0 < t - n\pi$

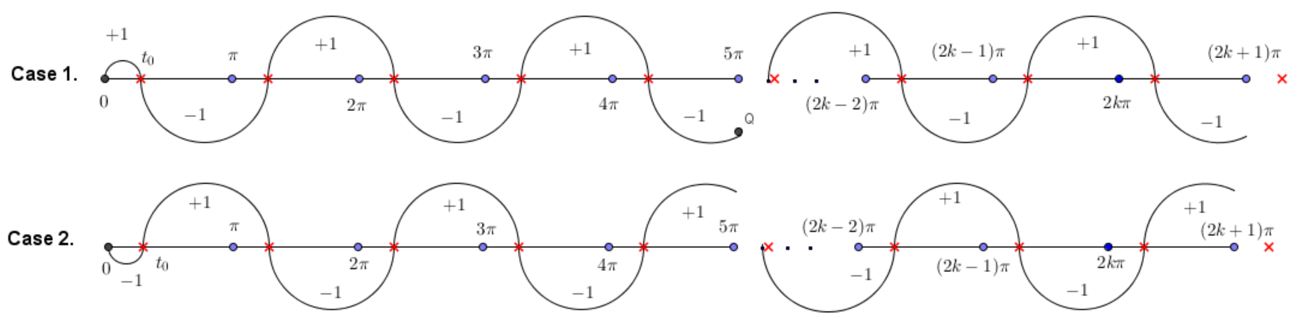


Figure 22: The controls in type 1.

Easily we obtain the formula for $z_0 = (x_0, v_0)$ in case 1 from (4.14)

$$(C1) \qquad z_0 = \begin{pmatrix} x_0 \\ v_0 \end{pmatrix} = \begin{pmatrix} -(2n+2)\cos t_0 + (-1)^n \cos t + 1 \\ -(2n+2)\sin t_0 + (-1)^n \sin t \end{pmatrix} \qquad t_0 \in [0, t - n\pi] \quad (4.15)$$

And since $(C2)$ and $(C1)$ are symmetric through 0, we obtain

$$(C2) \qquad z_0 = \begin{pmatrix} x_0 \\ v_0 \end{pmatrix} = \begin{pmatrix} (2n+2)\cos t_0 - (-1)^n \cos t - 1 \\ (2n+2)\sin t_0 - (-1)^n \sin t \end{pmatrix} \qquad t_0 \in [0, t - n\pi] \quad (4.16)$$

Type 2. u has n switching times. This type consist of two cases, which are symmetric through 0, with $t - n\pi < t_0 < \pi$

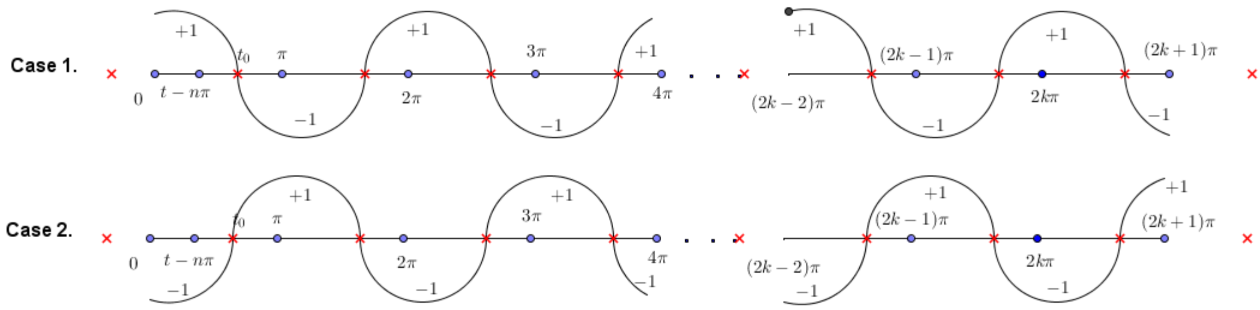


Figure 23: The controls in type 2.

Easily we obtain the formula for $z_0 = (x_0, v_0)$ in case 1 from (4.14)

$$(C3) \qquad z_0 = \begin{pmatrix} x_0 \\ v_0 \end{pmatrix} = \begin{pmatrix} -2n\cos t_0 - (-1)^n \cos t + 1 \\ -2n\sin t_0 - (-1)^n \sin t \end{pmatrix} \qquad t_0 \in [t - n\pi, \pi] \qquad (4.17)$$

And since $(C3)$ and $(C4)$ are symmetric through 0, we obtain

$$(C4) \qquad z_0 = \begin{pmatrix} x_0 \\ v_0 \end{pmatrix} = \begin{pmatrix} 2n\cos t_0 + (-1)^n \cos t - 1 \\ 2n\sin t_0 + (-1)^n \sin t \end{pmatrix} \qquad t_0 \in [t - n\pi, \pi] \qquad (4.18)$$

Note that at $t = n\pi$, the boundary of $\mathcal{R}(t)$ is always the circle center at $(0,0)$ with the radius of $2n$. We have some geometric presentations of $\mathcal{R}(t)$ when t is large.

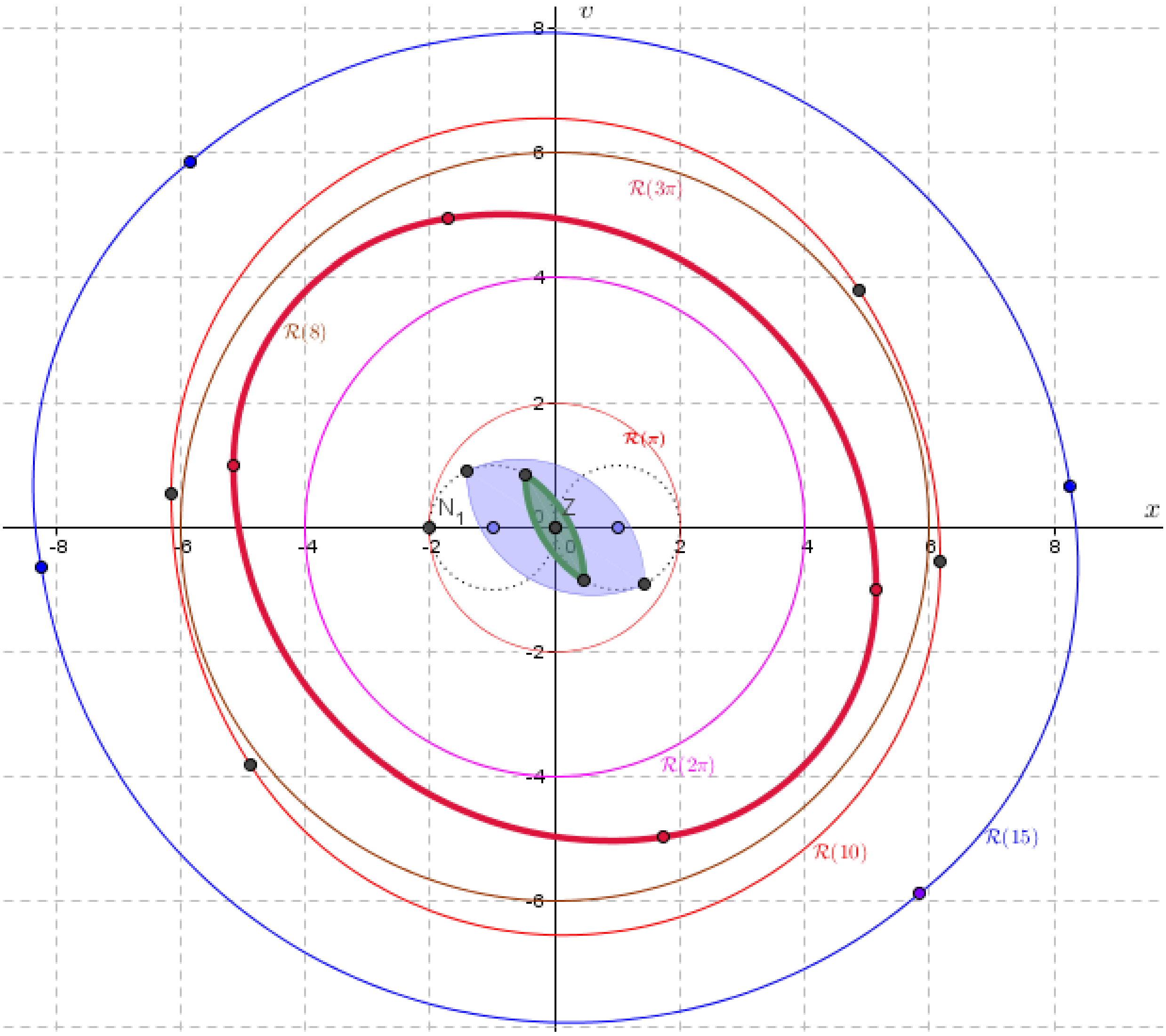


Figure 24: The reachable set $\mathcal{R}(t)$ when $t = 1, 2, \pi, 2\pi, 8, 3\pi, 10, 15$.

4.4.4 The minimum time function

Now from the formula of initial point which is determine the boundary of $\mathcal{R}(t)$, we deduce the formula for the minimum time function of this problem. We can estimate t and n in view of formula from reachable sets in general by consider the minimum $n \in \mathbb{N}$ such that

$$2n \leq \sqrt{x_0^2 + v_0^2} = \|z_0\| < 2(n+1)$$

It's implies $n\pi \leq t < (n+1)\pi$. First, consider case $0 < t \leq \pi$.

Case $0 < t \leq \pi$ A point $z_0 = (x_0, v_0) \in \partial\mathcal{R}(t)$ if and only if it can be presented in one of two forms

$$(Ca) \quad \begin{cases} x_0 = -2\cos t_0 + 1 + \cos t \\ v_0 = -2\sin t_0 + \sin t \end{cases} \qquad (Cb) \quad \begin{cases} x_0 = 2\cos t_0 - 1 - \cos t \\ v_0 = 2\sin t_0 - \sin t \end{cases}$$

which is respect to two kinds of optimal controls

$$(u_a) \qquad u_a(s) = \begin{cases} 1 & s \in [0, t_0] \\ 0 & s \in (t_0, t] \end{cases} \qquad\qquad (u_b) \qquad u_b(s) = \begin{cases} 0 & s \in [0, t_0] \\ 1 & s \in (t_0, t] \end{cases}$$

We derive the equations for t in both cases

$$\begin{array}{ll} (x_0 - (1+\cos t))^2 + (v_0 - \sin t)^2 = 4 & (x_0 + (1+\cos t))^2 + (v_0 + \sin t)^2 = 4 \\ (x_0, v_0) \in S\Big((1+\cos t, \sin t), 2\Big) & (x_0, v_0) \in S\Big((-1-\cos t, -\sin t), 2\Big) \\ ((1+\cos t), \sin t) \in S(z_0, 2) \cap S((1,0),1) & ((-1-\cos t), -\sin t) \in S(z_0, 2) \cap S((-1,0),1) \end{array}$$

where $S(x,r) = \{y \in \mathbb{R}^2 : \|x-y\| = r\}$ denote the circle center at x with radius r. Note that we have

$$\begin{aligned} S(z_0,2) \cap S((\ 1,0),1) \ &\neq \emptyset \Longleftrightarrow 1 \le \|z_0 - (\ 1,0)\| \le 3 \\ S(z_0,2) \cap S((-1,0),1) &\neq \emptyset \Longleftrightarrow 1 \le \|z_0 - (-1,0)\| \le 3 \end{aligned}$$

Since $z_0 \in \mathcal{R}(\pi)$, we have $\|z_0\| \le 2$. It's easy to see that there always exists at least one of two above equations has solution, from the triangle inequality.
In each cases, we need to find the intersection (x, y) of

$$(x,y) \in S(z_0,2) \cap S((\ 1,0),1) \qquad \text{such that } x \ge 0, y \ge 0 \tag{4.19}$$

$$(x,y) \in S(z_0,2) \cap S((-1,0),1) \qquad \text{such that } x \le 0, y \le 0 \tag{4.20}$$

We can easily see that at most one of both equation has a solution with this property. It's easy to see that the intersections of $(CA) : (xA, yA), rA$ and $(CB : (xB, yB), rB$ in general are given by (if its exist)

$$\begin{aligned} x_{12} &= \frac{1}{2}(x_B + x_A) + \frac{1}{2}\frac{(x_B - x_A)(r_A^2 - r_B^2)}{d^2} \pm 2(y_B - y_A)\frac{K}{d^2} \\ y_{12} &= \frac{1}{2}(y_B + y_A) + \frac{1}{2}\frac{(y_B - y_A)(r_A^2 - r_B^2)}{d^2} \pm -2(x_B - x_A)\frac{K}{d^2} \end{aligned}$$

where $d = \|(x_A, y_A) - (x_B, y_B)$ and

$$K = \frac{1}{4}\sqrt{\Big[(r_A + r_B)^2 - d^2\Big]\Big[d^2 - (r_A - r_B)^2\Big]}$$

Using this result, we can find all the intersection

$$(a) \qquad (x_1^a, y_1^a), (x_2^a, y_2^a) \qquad\qquad (b) \qquad (x_1^b, y_1^b), (x_2^b, y_2^b)$$

After choosing the intersection points has our's property, (after this step, we only have at most 2 solutions) we subtracting this to the equation of t

$$\begin{cases} 1 + \cos t & = x \\ \sin t & = y \end{cases} \qquad \text{or} \qquad \begin{cases} -1 - \cos t & = x \\ -\sin t & = y \end{cases}$$

respect to the property of the solution from (4.19) and (4.20). The explicit formula is very complicated to write here. Now, we can use any computer algebra system to find the graph of minimum time function in case $\|z_0\| \le 2$, i.e $0 < t \le \pi$.

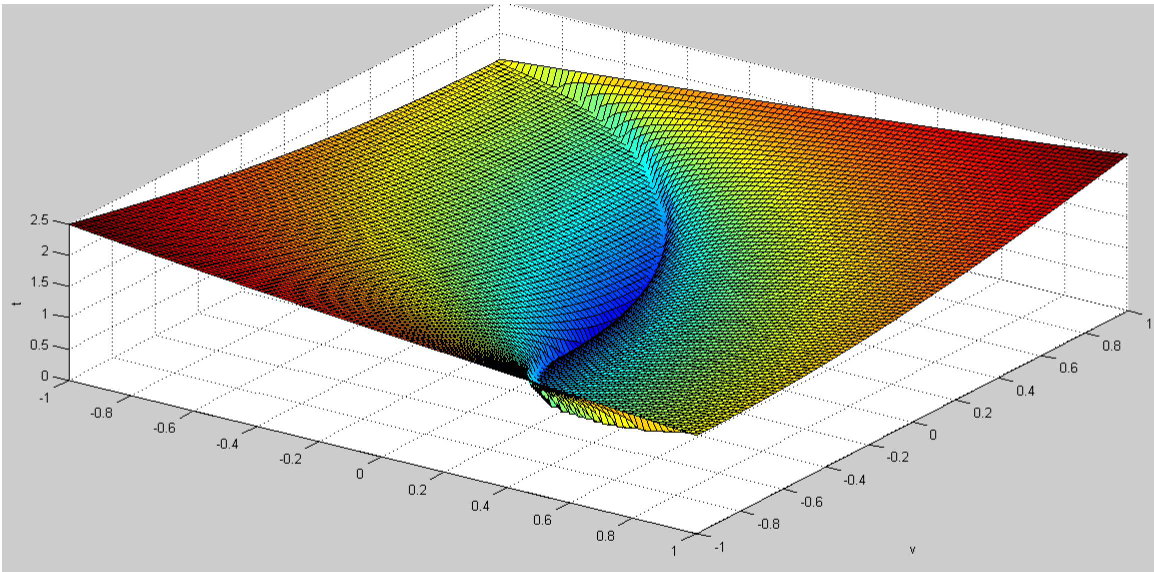


Figure 25: Graph of the minimum time function $T(z_0)$ when $\|z_0\| \leq 2$, i.e $T(z_0) \leq \pi$.

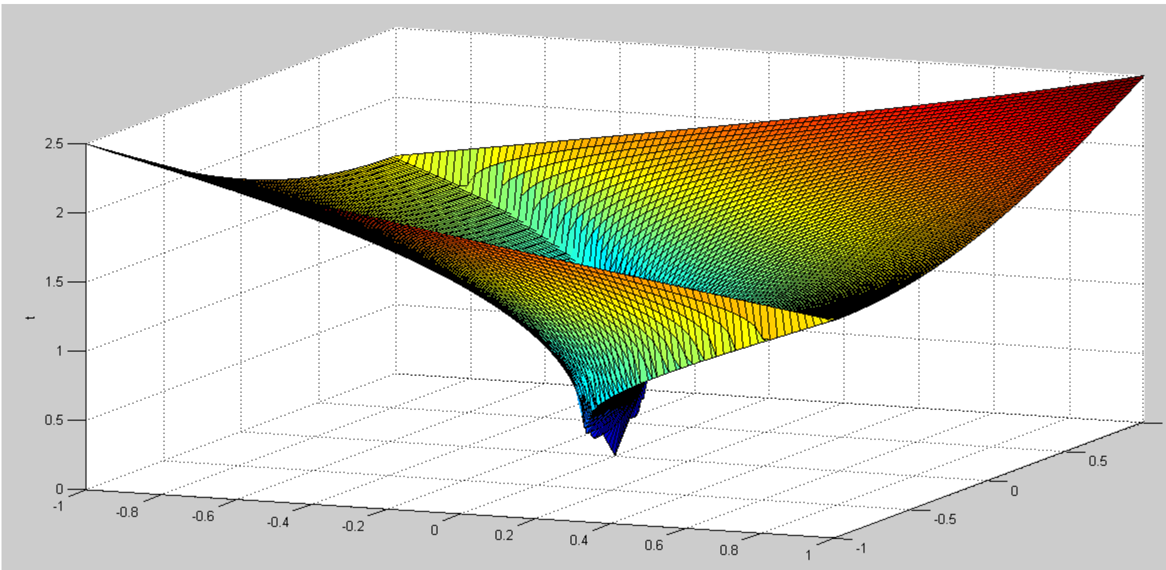


Figure 26: Graph of the minimum time function $T(z_0)$ when $\|z_0\| \leq 2$, i.e $T(z_0) \leq \pi$.

It's easy to see that this graph is not smooth in this case.

Case $\pi < t < \infty$ In this case, we have $\|z_0\| > 2$. The strategy to find t is determined by

1. Find the positive number $n \in \mathbb{N}$ such that

$$2n \leq \|z_0\| < 2n + 2 \Longrightarrow n = \left\lfloor \frac{\|z_0\|}{2} \right\rfloor$$

   Such an n like that is unique. Furthermore, this implies that $T(z_0) = t \in [n\pi, (n+1)\pi)$

2. From this and the formula of $\partial \mathcal{R}(t)$, (4.15), (4.16), (4.17), (4.18) we can see that t must have one of 4 forms

(C1) $(x, y) = ((-1)^n \cos t + 1, (-1)^n \sin t) \in S(z_0, 2n+2) \cap S((1,0), 1)$
where $x \geq 0, y \geq 0$

(C2) $(x, y) = (-(-1)^n \cos t - 1, -(-1)^n \sin t) \in S(z_0, 2n+2) \cap S((-1,0), 1)$
where $x \leq 0, y \leq 0$

(C3) $(x, y) = (-(-1)^n \cos t + 1, -(-1)^n \sin t) \in S(z_0, 2n) \cap S((1,0), 1)$
where $x \geq 0, y \leq 0$

(C4) $(x, y) = ((-1)^n \cos t - 1, (-1)^n \sin t) \in S(z_0, 2n) \cap S((-1,0), 1)$
where $x \leq 0, y \geq 0$

3. After having $\{(x_i, y_i)\}$ has this properties. We solve equations to find t.

4. Choosing t min.

To summarize this, similarly, the explicit formula is very complicated to write here, we will use a computer algebra system to write the graph of the minimum time function in this case.

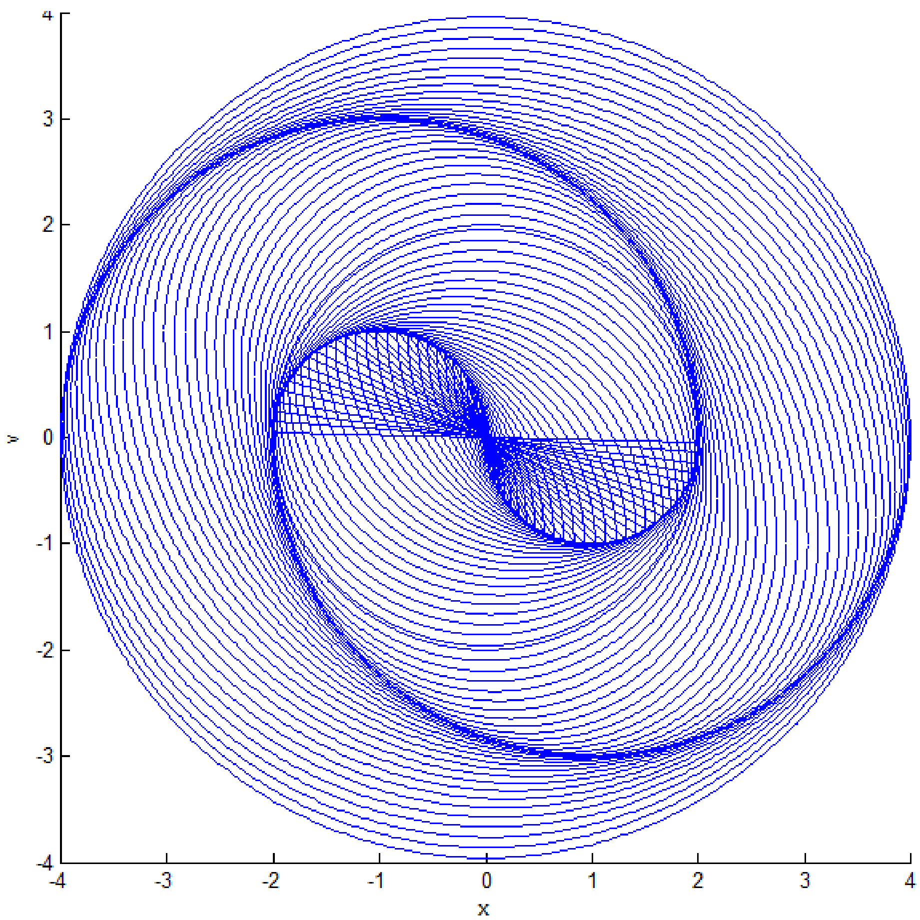


Figure 27: Graph of the minimum time function $T(z_0)$

We will see the graph of this minimum time function via some viewpoints

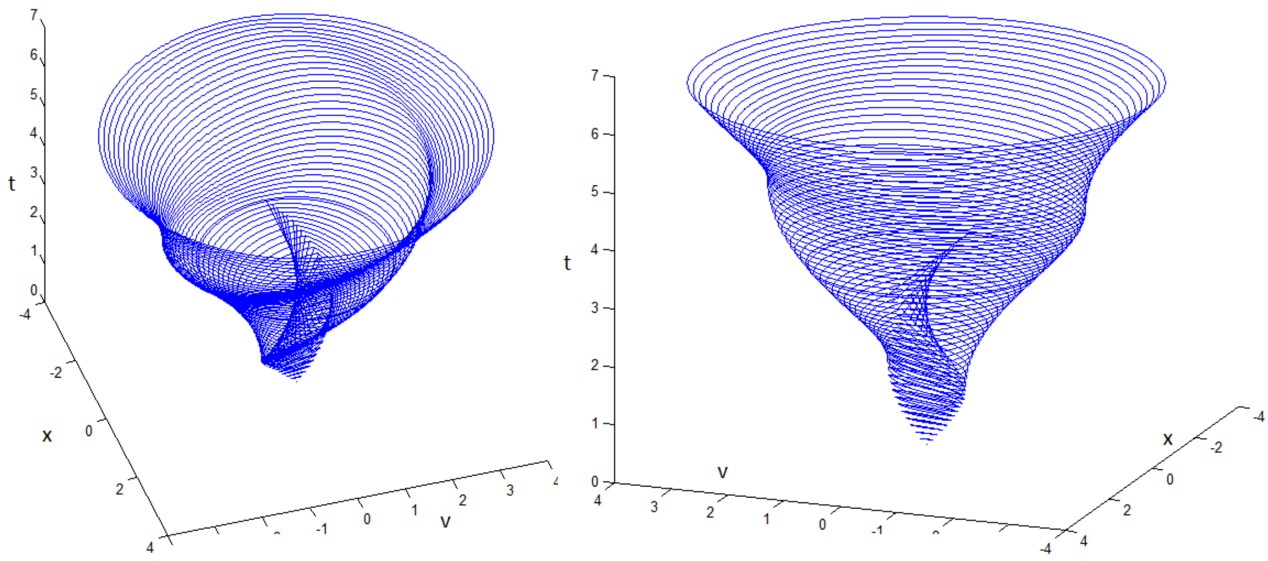


Figure 28: Graph of the minimum time function $T(z_0)$ via some viewpoints

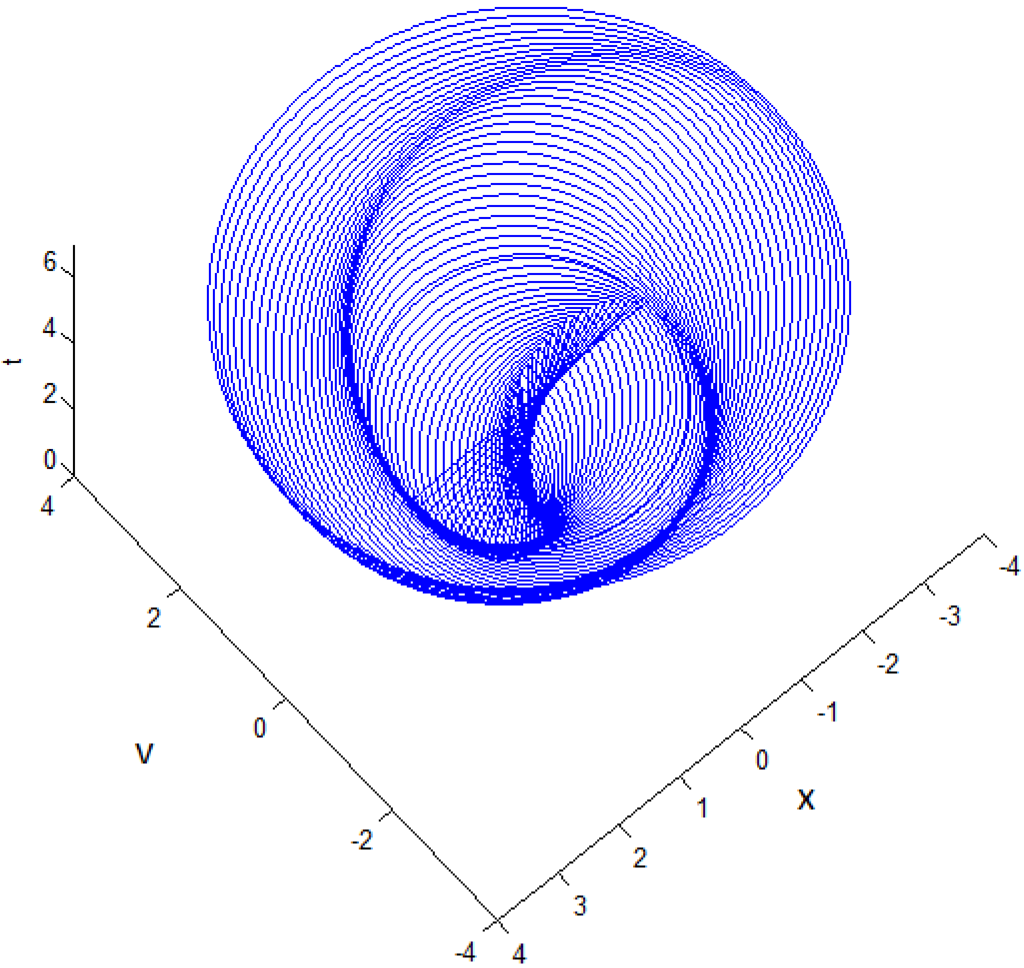


Figure 29: Graph of the minimum time function $T(z_0)$ via some viewpoints

From this, we see that $\partial\mathcal{R}(t)$ although it not smooth when $t \leq \pi$, but it's smooth for all $t \geq \pi$. It's little bit different to the rocket car problem. Also, we can see in general the graph of T is not smooth in whole space $\mathbb{R}^n$. This inspire us to study the regularity of the minimum time function. Before doing this, we will prove a tool which is mainly used later.

4.5 Another way to rewrite the Pontryagin maximum principle

Consider the linear control system where $A \in \mathbb{M}^{n\times n}, B \in \mathbb{M}^{n\times m}$ and $u : [0,+\infty) \longrightarrow \mathbb{R}^m$,[2]

$$\begin{cases} x'(t) &= Ax(t) + Bu(t) \\ x(0) &= x_0 \end{cases} \tag{4.21}$$

Let's denote the optimal control in time $T = T(x_0)$ is $u^*(t)$, and let the trajectory corresponding is

$$\mathbf{x}(t) = y^{x_0,u^*}(t) \qquad \forall\ t \in [0,T]$$

Recall from theorem 4.7, there exists $0 \neq \theta \in \mathbb{R}^n$ such that

$$\langle \theta, e^{-tA}Bu^*(t)\rangle = \max_{a\in U}\langle \theta, e^{-tA}Ba\rangle \qquad \Longleftrightarrow \qquad \theta^T e^{-tA}Bu^*(t) = \max_{a\in U}\left\{\theta^T e^{-tA}Ba\right\}$$

Introduce the new operator, called the Hamiltonian

$$\overline{H}(x,p,a) = \langle Ax,p\rangle + \langle Ba,p\rangle = p^T(Ax+Ba)$$

We can re-write the Potryagin maximum principle 4.7 as following

Theorem 4.10: Let u^* be the optimal control in time $T = T(x_0)$ and $x(.)$ is the corresponding trajectory of the system (4.21), then there exists a absolutely continuous function $\lambda : [0,T] \longrightarrow \mathbb{R}^n$, is the costate function which is never vanish such that

- $\mathbf{x}(.)$ satisfies the ODE
$$\mathbf{x}'(t) = \nabla_p\overline{H}(\mathbf{x}(t),\lambda(t),u^*(t)) \tag{ODE}$$
- $\lambda(.)$ satisfies the adjoint equations ADJ
$$\lambda'(t) = -\nabla_x\overline{H}(\mathbf{x}(t),\lambda(t),u^*(t)) \tag{ADJ}$$
- We have the maximization principle
$$\overline{H}(\mathbf{x}(t),\lambda(t),u^*(t)) = \max_{a\in U}\overline{H}(\mathbf{x}(t),\lambda(t),a) \tag{M}$$

Proof. Let θ in the theorem 4.7 and consider the control system

$$\begin{cases} \lambda'(t) &= -A^T\lambda(t) \\ \lambda(0) &= \theta \end{cases} \tag{4.22}$$

[2]Note that in this case we understand $\mathbb{R}^m$ as $\mathbb{M}^{m\times 1}(\mathbb{R})$

Now let the solution is $\lambda(t) = e^{-tA^T}\theta$, so $\left(\lambda(t)\right)^T = \theta^T e^{-tA}$, now clearly from theorem 4.7 we get

$$\theta^T e^{-tA} B u^*(t) = \max_{a\in U}\left\{\theta^T e^{-tA} B a\right\}$$

So this implies that

$$\lambda(t)^T B u^*(t) = \max_{a\in U}\lambda(t)^T Ba \quad \Longrightarrow \quad \lambda(t)^T\left(A\mathbf{x}(t) + Bu^*(t)\right) = \max_{a\in U}\lambda(t)^T\left(A\mathbf{x}(t) + Ba\right)$$

i.e,

$$\overline{H}(\mathbf{x}(t), \lambda(t), u^*(t)) = \max_{a\in U}\overline{H}(\mathbf{x}(t), \lambda(t), a)$$

therefore (M) is true. Now (ODE) is also true since

$$\nabla_p\overline{H}(x,p,a) = \frac{\partial}{\partial p}\left[p\cdot(Ax+Ba)\right] = Ax + Ba \Longrightarrow \nabla_p\overline{H}(\mathbf{x}(t), \lambda(t), u^*(t)) = \mathbf{x}'(t)$$

Finally, we have

$$\nabla_x\overline{H}(x,p,a) = \begin{pmatrix} \frac{\partial H}{\partial x_1} \\ \frac{\partial H}{\partial x_2} \end{pmatrix}(x,p,a) = A^T\cdot p \Longrightarrow \nabla_x\overline{H}(\mathbf{x}(t), \lambda(t), u^*(t)) = A^T\lambda(t) = -\lambda'(t)$$

Thus the proof is complete, because the absolutely continuous property of λ is trivial. □

Similarly to above theorem, we introduce the minimization version of Potryagin Maximum principle as follow, observe that if x_0 is steered to 0 by the control u^*, which is respect to the trajectory $\mathbf{x}(.) = y^{x_0,u^*}(.)$. Then 0 can be steered into x_0 in time T by the reversed control $v^*(t) = u^*(T-t)$, in this case the corresponding trajectory is

$$\mathbf{z}(t) = y^{x_0,u^*}(T-t) = \mathbf{x}(T-t)$$

Also define the minimized Hamiltonian

$$H(x,p) = \min_{a\in U}\left[p\cdot(Ax+Ba)\right] = \min_{a\in U}\overline{H}(x,p,a)$$

Theorem 4.11: Let u^* be the optimal control in time $T = T(x_0)$ and $x(.)$ is the corresponding trajectory of the system (4.21), then there exists a absolutely continuous function $\psi : [0,T] \longrightarrow \mathbb{R}^n$, is the which is never vanish such that

(i) $\psi'(t) = \psi(t)\cdot A^T$.

(ii) $\overline{H}(\mathbf{x}(T-t), \psi(t), u^*(T-t)) = H(\mathbf{x}(T-t), \psi(t))$ for a.e $t\in[0,T]$.

(iii) The mapping $t \longmapsto H(\mathbf{x}(T-t), \psi(t))$ is constant for all $t\in[0,T]$.

(iv) The vector $\psi(T)$ belong to the normal cone of $\mathcal{R}(T)$ at x_0, i.e $\psi(T)\in N_{\mathcal{R}(T)}(x_0)$.

(v) For each $t\in(0,T)$, we also have the vector $\psi(t)$ belong to the normal cone of $\mathcal{R}(t)$ at $\mathbf{x}(T-t)$, i.e $\psi(t)\in N_{\mathcal{R}(t)}(\mathbf{x}(T-t))$.

Proof. Consider the following function, where λ is the costate function in theorem 4.10

$$\begin{aligned} \psi : [0,T] &\longrightarrow \mathbb{R}^n \\ t &\longmapsto -\lambda(T-t) = -e^{-(T-t)A^T}\theta \end{aligned}$$

We have $\psi'(t) = \lambda'(T-t)$, now from above theorem, we have

$$\lambda(T-t)^T Bu^*(T-t) = \max_{a\in U}\left\{\lambda(T-t)^T Ba\right\} \implies -\lambda(T-t)^T Bu^*(T-t) = \min_{a\in U}\left\{-\lambda(T-t)^T Ba\right\}$$

i.e,

$$\psi(t)^T Bu^*(T-t) = \min_{a\in U}\psi(t)^T Ba \implies \psi(t)^T\cdot\Big(A\mathbf{x}(T-t)+Bu^*(T-t)\Big) = \min_{a\in U}\psi(t)\cdot\Big(A\mathbf{x}(T-t)+Ba\Big)$$

In otherword, we have

$$\overline{H}(\mathbf{x}(T-t),\psi(t),u^*(T-t)) = \min_{a\in U}\overline{H}(\mathbf{x}(T-t),\psi(t),a) = H(\mathbf{x}(T-t),\psi(t))$$

So (i) and (ii) is true. For (iii) by (ADJ) in above theorem, we have

$$\begin{aligned} \psi'(t) = \lambda'(T-t) &= -\nabla_x\overline{H}(\mathbf{x}(T-t),\lambda(T-t),u^*(T-t)) \\ &= \nabla_x\overline{H}(\mathbf{x}(T-t),\psi(t),u^*(T-t)) = \nabla_x H(\mathbf{x}(T-t),\psi(t)) \end{aligned}$$

So by (ODE) in above theorem, we have

$$\begin{aligned} \frac{\partial}{\partial t}H(\mathbf{x}(T-t),\psi(t)) &= -\underbrace{\nabla_x H(\mathbf{x}(T-t),\psi(t))}_{\psi'(t)}\mathbf{x}'(T-t) + \underbrace{\nabla_p H(\mathbf{x}(T-t),\psi(t))}_{\mathbf{x}'(T-t)}\psi'(t) \\ &= -\psi'(t)\mathbf{x}'(T-t) + \mathbf{x}'(T-t)\psi'(t) = 0 \end{aligned}$$

So $t \longmapsto H(x(T-t),\psi(t))$ is constant. Finally we have $\psi(T) = -\lambda(0) = -\theta = \xi$ where ξ is chosen in theorem 4.7, which is clearly the normal vector by our construction

$$\langle \xi, y - x_0\rangle \le 0 \qquad \forall \quad y \in \mathcal{R}(t)$$

So we have $\psi(T) \in N_{\mathcal{R}(t)}(x_0)$.

(v) By the principle of optimality 2.69, for each $t \in (0,t)$ we have

$$T(\mathbf{x}(T-t)) = t \Longrightarrow \mathbf{x}(T-t) \in \partial\mathcal{R}(T)$$

Doing similarly as the principle of optimality 2.69, we can find the corresponding optimal control v of $\mathbf{x}(T-t)$ very fast. But in here we will find it by the optimal trajectory of $\mathbf{x}(0)$ to see this fact more deeply. For $s \in [0,t]$ we have

$$\begin{aligned} \mathbf{x}(T-t+s) &= e^{(T-t+s)A}x_0 + \int_0^{T-t+s} e^{(T-t+s-\delta)A}Bu^*(\delta)\,d\delta \\ &= e^{sA}e^{(T-t)A}x_0 + \int_0^{T-t} e^{(T-t+s-\delta)A}Bu^*(\delta)\,d\delta + \int_{T-t}^{T-t+s} e^{(T-t+s-\delta)A}Bu^*(\delta)\,d\delta \\ &= e^{sA}\left(e^{(T-t)A}x_0 + \int_0^{T-t} e^{(T-t-\delta)A}Bu^*(\delta)\,d\delta\right) + \int_0^s e^{(s-\eta)A}Bu^*(\eta+T-t)\,d\eta \\ &= e^{sA}\mathbf{x}(T-t) + \int_0^s e^{(s-\eta)A}Bu^*(\eta+T-t)\,d\eta \end{aligned}$$

So if we define the new control

$$v(\eta) = \begin{cases} u^*(\eta + T - t) & 0 \le \eta \le t \\ 0 & \text{otherwise} \end{cases}$$

Thus, we have the optimal control v steer $\mathbf{x}(T-t)$ to the origin has the form

$$v(\eta) = \begin{pmatrix} v_1(\eta) \\ \vdots \\ v_m(\eta) \end{pmatrix} \qquad \text{where} \qquad v_i(\eta) = u_i^*(\eta + T - t) = -\operatorname{sign}\langle \zeta, e^{-(\eta+T-t)A} b_i \rangle$$

Recall that $\psi(s) = \langle \zeta, e^{-(T-s)A} \rangle$, so we obtain

$$v_i(\eta) = -\operatorname{sign}\langle \zeta, e^{(T-t)A} e^{-\eta A} b_i \rangle = -\operatorname{sign}\langle \psi(t), e^{-\eta A} b_i \rangle$$

Now by theorem 4.9 we obtain $\psi(t) \in N_{\mathcal{R}(t)}(\mathbf{x}(T-t))$. $\square$

Theorem 2.69 give us an image how we can reduce a initial point $x_0 \in \partial\mathcal{R}(T)$ into a point x' also live in the trajectory $x_0 \longrightarrow 0$ with $x' \in \partial\mathcal{R}(t)$ for $t < T$, and

$$x' = y^{x_0,u^*}(T-t) \qquad \text{and} \qquad T(x') = T(y^{x_0,u^*}(T-t)) = t$$

Similarly, we now establish something look similar to this, but expand outside $\mathcal{R}(T)$. The main tool used to prove is the structure of the optimal control, given by the maximum principle above and the dynamic reverse system.

Definition 4.12: For $0 \neq x \in \mathcal{R}$, we call x is an **optimal point** if for every $t > T(x)$, we can find some $x_t \in \partial\mathcal{R}(t)$ and some control u_t such that

$$y^{x_t,u_t}(t - T(x)) = x$$

Now using theorem 4.9, we will show that any non-zero point in $\mathcal{R}$ is optimal

Theorem 4.13 (Every point is optimal)**:** If $x_\alpha \in \partial\mathcal{R}(\alpha)$, then for every $\beta > \alpha$, there exists $x_\beta \in \partial\mathcal{R}(\beta)$ such that the optimal trajectory of x_β, which is respect to the optimal control u_β, has the property

$$y^{x_\beta,u_\beta}(\beta - \alpha) = x_\alpha \tag{4.23}$$

Proof. Pick $\zeta \in N_{\mathcal{R}(t)}(x_\alpha)$. By theorem 4.10 and 4.11, setting $\theta = -\zeta$, the following functions

$$\begin{aligned} \lambda : [0,\alpha] &\longrightarrow \mathbb{R}^n \\ t &\longmapsto \langle \theta, e^{-tA} \rangle \end{aligned} \qquad\qquad \begin{aligned} \psi : [0,\alpha] &\longrightarrow \mathbb{R}^n \\ t &\longmapsto -\lambda(\alpha - t) = \left\langle \zeta, e^{(\alpha-t)A} \right\rangle \end{aligned}$$

is well–defined and has the property $\psi(\alpha) = \zeta$. We defined the extension of ψ by $\Psi : [0,\beta] \longrightarrow \mathbb{R}^n$ by natural way

$$\begin{aligned} \Psi : [0,\beta] &\longrightarrow \mathbb{R}^n \\ t &\longrightarrow \left\langle \zeta, e^{(\alpha-t)A} \right\rangle \end{aligned}$$

One can see that Ψ is well-defined, and furthermore, for $0 \le t \le \alpha$ then

$$\Psi(t) \equiv \psi(t) \qquad \text{i.e} \qquad \Psi|_{[0,\alpha]} = \psi$$

We can image that ψ is the adjoint function send each t into the normal vector $\psi(t)$ of $\mathcal{R}(t)$ of point $y^{x_\alpha,u_\alpha}(T-t)$, this fact is the consequence of theorem 4.11 and the principle of optimality 2.69. Now by Pontryagin's maximum principle, the control u_α which steers x_α to the origin must have form

$$u_\alpha(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_m(t) \end{pmatrix} \qquad \text{where} \qquad u_i(t) = -\operatorname{sign}\langle \zeta, e^{-tA} b_i\rangle = -\operatorname{sign}\langle \psi(\alpha), e^{-tA} b_i\rangle$$

for all $t \in [0,\alpha]$. This suggest that we can establish the new control u_β which steers x_β to x_α by

$$u_\beta(t) = \begin{pmatrix} v_1(t) \\ \vdots \\ v_m(t) \end{pmatrix} \qquad \text{where} \qquad v_i(t) = -\operatorname{sign}\langle \Psi(\beta), e^{-tA} b_i\rangle$$

for all $0 \le t \le \beta - \alpha$, where $b_1, b_2, \dots, b_m$ are columns of B. Now consider

$$x_\beta = -\int_0^\beta e^{-tA} B u_\beta(t)\, dt \in \mathcal{R}(\beta) \tag{4.24}$$

This construction is reasonable because as we already discuss above, if everything in our theorem is true, then

$$\Psi(t) \in N_{\mathcal{R}(t)}(y^{x_\beta,u_\beta}(\beta - t))$$

This is compatible with our consumption at $t = \alpha$, since

$$\zeta = \Psi(\alpha) = \psi(\alpha) \in N_{\mathcal{R}(\alpha)}(y^{x_\beta,u_\beta}(\beta-\alpha)) = N_{\mathcal{R}(\alpha)}(x_\alpha)$$

So all the rest is check that $x \in \partial\mathcal{R}(\beta)$ and the condition (4.23) is true.

(i) Check $x \in \partial\mathcal{R}(\beta)$, this is simply the consequence of theorem 4.9. Since $\Psi(\beta) \in H(A,B)$ is clear and the formula (4.24) give us $x_\beta \in \mathcal{R}(\beta)$. So by theorem 4.9 we have $x \in \partial\mathcal{R}(\beta)$ and $\Psi(\beta) \in \mathcal{R}(\beta)$.

(ii) Check the condition (4.23) is true , i.e $y^{x_\beta,u_\beta}(\beta-\alpha) = x_\alpha$. It's just a simple calculation from the formula

$$y^{x_\beta,u_\beta}(t) = e^{tA}x_\beta + \int_0^t e^{(t-s)A} B u_\beta(s)\, ds \tag{4.25}$$

and the simple observe that for all $t \in [0,\alpha]$ we have $\Psi(\beta) = \left\langle \zeta, e^{(\alpha-\beta)A} b_i \right\rangle$, so

$$\begin{aligned} u_i^\beta(t+\beta-\alpha) &= -\operatorname{sign}\left\langle \Psi(\beta), e^{-(t+\beta-\alpha)A} b_i \right\rangle \\ &= -\operatorname{sign}\left\langle \zeta, e^{(\alpha-\beta+t+\beta-\alpha)A} b_i \right\rangle = -\operatorname{sign}\left\langle \zeta, e^{-tA} b_i \right\rangle = u_i^\alpha(t) \end{aligned}$$

This is true for all $i = 1, 2, \dots, m$, so

$$u_\beta(t + (\beta - \alpha)) = u_\alpha(t) \qquad \forall\, 0 \le t \le \alpha \tag{4.26}$$

Now combine (4.24) and (4.25) we have

$$\begin{aligned}
y^{x_\beta,u_\beta}(\beta-\alpha) &= -e^{(\beta-\alpha)A}\int_0^\beta e^{-sA}Bu_\beta(s)\,ds + \int_0^{\beta-\alpha} e^{(\beta-\alpha-s)A}Bu_\beta(s)\,ds \\
&= -\int_0^\beta e^{(\beta-\alpha-s)A}Bu_\beta(s)\,ds + \int_0^{\beta-\alpha} e^{(\beta-\alpha-s)A}Bu_\beta(s)\,ds \\
&= -\int_{\beta-\alpha}^\beta e^{(\beta-\alpha-s)A}Bu_\beta(s)\,ds \\
&= -\int_0^\alpha e^{(\alpha-t)A}Bu_\beta(t+(\beta-\alpha))\,dt \\
&= -\int_0^\alpha e^{-tA}Bu_\alpha(s)\,dt = x_\alpha
\end{aligned}$$

by changing variable $s = t + (\beta - \alpha)$ and using (4.26).

Thus, the proof is complete. □

5 The regularity of minimum time function

As we discuss through two former examples, we see that in general the minimum time function is not smooth in the whole domain $\mathcal{R}$, indeed, recall the rocket road car problem, we will prove that the minimum time function is even non-Lipschitz.

As we already prove in 4.3.2, that is the boundary $\mathcal{R}(t)$ consist of two curves, one of them have the following form

$$\begin{cases} x_0 & = t_0^2 - \frac{t^2}{2} \\ v_0 & = t - 2t_0 \end{cases}$$

where t_0 is the point in $[0,1]$ such that the corresponding optimal control u_a of (x_0, v_0) is

$$u_a(s) = \begin{cases} 1 & \text{on } [0, t_0] \\ 0 & \text{on } (t_0, 1] \end{cases}$$

Now assume T is Lipschitz in $\mathcal{R} = \mathbb{R}^2$ in this case, then there exists a constant $C > 0$ such that

$$|T(x) - T(y)| \leq C\|x - y\| \qquad \forall\, x, y \in \mathcal{R} \qquad \Longrightarrow \qquad T(x) \leq C\|x\| \qquad \forall\, x \in \mathcal{R}$$

Now for $t > 0$, we choose the points $(x_0^t, y_0^t) \in \mathcal{R}(n)$ which lie on the x-axis, i.e $v_0^t = 0$

$$v_0^t = t - 2t_0 = 0 \Longrightarrow t_0 = \frac{t}{2} \Longrightarrow x_0^t = \frac{t^2}{4} - \frac{t^2}{2} = -\frac{t^2}{4}$$

Using sequence $(x_0^t, v_0^t) = \left(-\frac{t^2}{4}, 0\right) \in \partial\mathcal{R}(t)$, we have

$$t = T(x_0^t, v_0^t) \leq C \left\| \left(-\frac{t^2}{4}, 0\right) \right\| = \frac{Ct^2}{4} \Longrightarrow \frac{4}{C} \leq t$$

for all $t > 0$, let $t \longrightarrow 0$ we get a contradiction.

We want to study the regularity of the minimum time function T, with the following assumption to ensure that $T : \mathbb{R}^n \longrightarrow \mathbb{R}$ is continuous by theorem 3.11

Assumption

$$\begin{cases} \text{rank } G(A, B) = n \text{ (Rank-Kalman's condition)} \\ \text{Re } \lambda \leq 0 \text{ for each eigenvalue } \lambda \text{ of } A \end{cases} \tag{$\star$}$$

In this case $T : \mathbb{R}^n \longrightarrow \mathbb{R}$. Our work is motivate by the following theorem in [3]

Theorem 5.1 (Giovanni Colombo, Antonio Marigonda, Peter R.Wolenski): Assume the linear control system (2.25) has the target S be convex, and T is continuous on $\overline{\mathbb{R}^n \setminus S}$, then the epigraph epi$(T)$ has positive reach.

From this, in our case $S = \{0,\}$ and under $(\star)$, we obtain

Theorem 5.2: Assume the linear control system (2.25) satisfies ($\star$), then

- The minimum time- function T is continuous from $\mathcal{R} = \mathbb{R}^n \longrightarrow \mathbb{R}$.
- The epigraph $\text{epi}(T)$ has positive reach.

Theorem 5.3 (Giovanni Colombo, Khai T. Nguyen, Antonio Marigonda): Given $f : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}$ is continuous and Ω is open, assume

- f is continuous.
- The epigraph $\text{epi}(f)$ has positive reach

then the following set

$$S_f^\infty = \{x : f \text{ is not Lipschitz at } x\}$$

is satisfied

(i) S_f^∞ is closed in Ω.

(ii) $\mathcal{L}^n(S_f^\infty) = 0$, where $\mathcal{L}^n$ is the Lebesgue measure on $\mathbb{R}^n$.

Under the assumption ($\star$), we have $\mathbb{R}^n \backslash S_f^\infty$ is open, $\mathcal{L}^n(S_f^\infty) = 0$, and

- $T : \mathbb{R}^n \backslash S_f^\infty$ is locally Lipschitz
- The epigraph $\text{epi}(T)$ has positive reach.
- $\mathcal{L}^n(S_T^\infty) = 0$.

So by theorems 2.55 and 2.56, we obtain the following corollary

Corollary 5.4: Under the assumption ($\star$), we have $T : \mathbb{R}^n \backslash S_f^\infty \longrightarrow \mathbb{R}$ is locally semi-convex. And as a consequence from the H.Rademacher's theorem 2.9, we have $T : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a.e twice differentiable.

This fact match perfectly with two example we already discussed former. So the main goal of us is now study the set S_T^∞.

Main goal. Study the set S_T^∞.

In this section, we always assume the assumption ($\star$) to ensure $\mathcal{R} = \mathbb{R}^n$.

5.1 Viscosity solutions for Linear control systems

Because the viscosity solution is related with the differentiability of the solution. We will establish the Hamilton-Jacobi equation which has T as a viscosity solution.

Theorem 5.5: Define the minimized Hamiltonian

$$H(x, p) := \min_{w \in U} \Big[p \cdot (Ax + Bw) \Big] \qquad \forall\, x \in \mathcal{R}$$

then T is a viscosity solution of Hamilton–Jacobi–Bellman equation

$$-H(x, \nabla T(x)) - 1 = 0 \qquad \forall\, x \in \mathcal{R}\backslash\{0\}$$

i.e,T is a viscosity solution of Hamilton–Jacobi–Bellman equation

$$-\min_{w\in U}\Big[p\cdot(Ax+Bw)\Big] - 1 = 0 \qquad \forall\, x \in \mathcal{R}\backslash\{0\} \tag{5.1}$$

Proof. Let $\varphi \in C^1(\mathcal{R}\backslash\{0\})$.

- **Sub-solution**. If $T-\varphi$ attains a local maximum at a point $x_0 \in \mathcal{R}\backslash\{0\}$, we will prove

$$-H(x, \nabla\varphi(x)) - 1 \le 0 \Longleftrightarrow \min_{w\in U}\Big(\nabla\varphi(x_0)\cdot(Ax_0+Bw)\Big) + 1 \ge 0 \tag{5.2}$$

If (5.2) is not true, then there exists $w \in U$ and $\theta > 0$ such that

$$\nabla\varphi(x_0)\cdot(Ax_0+Bw) + 1 < -2\theta$$

Now let $x(.) = y^{x_0,w}(.)$ be the unique solution of the linear system

$$\begin{cases} x'(t) &= (Ax_0+Bw) \qquad t\in[0,+\infty) \\ x(0) &= x_0 \end{cases}$$

Since $Ax+Bw$ is bounded, and by continuity, there exists $\alpha > 0$ such that

$$\nabla\varphi(x)\cdot(Ax+Bw) + 1 < -\theta$$

wherever

$$|x-x_0| \le \alpha$$

Since $x(.) = y^{x_0,u}(.)$ is continuous at 0, there exists $s > 0$ such that

$$t < s \Longrightarrow |x(t)-x(0)| < \alpha$$

Since $T-\varphi$ attains local maximum at x_0, we must have

$$\begin{aligned} T(x(s)) - T(x(0)) \le \varphi(x(s)) - \varphi(x(0)) &= \int_0^s \frac{d}{dt}\varphi(x(t))\,dt \\ &= \int_0^s \nabla\varphi(x(t))\cdot x'(t)\,dt \\ &= \int_0^s \nabla\varphi(x(t))\cdot(Ax(t)+Bw)\,dt \\ &\le \int_0^s -(1+\theta)\,dt = -s(1+\theta) < -s \end{aligned}$$

But on the other hand, from the dynamic programming principle 2.68, we also have

$$T(x(0)) = T(x_0) \le s + T(y^{x_0,w}(s)) = s + T(x(s)) \Longrightarrow -s < T(x(s)) - T(x(0)) < -s$$

so this yields a contradiction, therefore (5.2) must be true, or T is a sub-solution of

$$-H(x, \nabla T(x)) - 1 = 0 \qquad \forall\, x \in \mathcal{R}\backslash\{0\}$$

- **Super-solution**. If $T-\varphi$ attains a local minimum at a point $x_0 \in \mathcal{R}\backslash\{0\}$, we will prove

$$-H(x,\nabla\varphi(x))-1 \geq 0 \Longleftrightarrow \min_{w\in U}\Big(\nabla\varphi(x_0)\cdot(Ax_0+Bw)\Big)+1 \leq 0 \tag{5.3}$$

If (5.3) is not true, then there exists $\theta > 0$ such that

$$\nabla\varphi(x_0)\cdot f(x_0,w)+1 > 2\theta \qquad \forall\, w \in U$$

Again, by boundedness of f and continuity, there exists $\alpha > 0$ such that

$$\nabla\varphi(x)\cdot(Ax+Bw)+1 > \theta \qquad \forall\, w \in U$$

wherever

$$|x-x_0| \leq \alpha$$

Since $x(.) = y^{x_0,u}(.)$ is continuous at 0, there exists $s > 0$ such that

$$t < s \Longrightarrow |x(t)-x(0)| < \alpha$$

Pick $u \in U_{ad}$ is a arbitrary control, and let $y^{x_0,u}(.) = x(.)$. Since $T-\varphi$ attains local minimum at x_0, we must have

$$\begin{aligned}
T(x(s))-T(x(0)) \geq \varphi(x(s))-\varphi(x(0)) &= \int_0^s \frac{d}{dt}\varphi(x(t))\,dt \\
&= \int_0^s \nabla\varphi(x(t))\cdot x'(t)\,dt \\
&= \int_0^s \nabla\varphi(x(t))\cdot(Ax(t)+Bu(t))\,dt \\
&\geq \int_0^s (-1+\theta)\,dt = s(-1+\theta)
\end{aligned}$$

Thus, it implies that

$$T(x_0)+s\theta < T(x(s))+s \Longleftrightarrow T(x_0)+s\theta < s+T(y^{x_0,u}(s)) \qquad \forall\, u \in U_{ad}$$

Taking the infimum both sides over all control $u \in \mathcal{U}_{ad}$, we see that

$$T(x_0)+s\theta \leq \inf_{u\in\mathcal{U}_{ad}}\Big(s+T(y^{x_0,u}(s))\Big) = T(x_0)$$

according to the dynamic programming principle 2.68. Thus this yields a contradiction, therefore (5.3) must be true, or T is a super-solution of

$$-H(x,\nabla T(x))-1 = 0 \qquad \forall\, x \in \mathcal{R}\backslash\{0\}$$

□

From this fact, we can see that if T is differentiable at $x_0 \in \partial\mathcal{R}(T)$, then we must have

$$H(x_0, \nabla T(x_0)) = -1 = \max_{a\in U}\langle \nabla T(x_0), Ax_0 + Ba\rangle \quad \Longrightarrow \quad \nabla T(x_0) \neq 0$$

In case $\partial\mathcal{R}(T)$ is smooth[3], then clearly $\nabla T(x_0)$ is the normal vector of $\partial\mathcal{R}(T)$ at x_0. So this suggest us to study $H(x_0, \zeta)$ where ζ the normal vector of $\partial\mathcal{R}(T)$ at x_0 in general case.

Theorem 5.6: Let $T = T(x_0)$ and $\zeta \in N_{\mathcal{R}(T)}(x_0)$, then

$$H(x_0, \zeta) \leq 0$$

Proof. Since $\zeta \in N_{\mathcal{R}(T)}(x_0)$ and $\mathcal{R}(T)$ is convex and closed, we have

$$\langle \zeta, y - x\rangle \leq 0 \qquad \forall\, y \in \mathcal{R}(T)$$

Let $u^*(.)$ is an optimal control and $\mathbf{x}(.) = y^{x_0,u^*}(.)$ be the corresponding trajectory, Now because $\mathbf{x}(t) \in \mathcal{R}(T)$ for all $t \in [0, T]$ we have

$$\langle \zeta, \mathbf{x}(t) - x_0\rangle = \langle \zeta, \mathbf{x}(t) - \mathbf{x}(0)\rangle\rangle \leq 0 \qquad \forall\, t \in [0, T] \tag{5.4}$$

By theorem 2.71, there exists $M = M(T)$ such that $\|\mathbf{x}(t) - \mathbf{x}(0)\| \leq Mt$ for all $t \in [0, T]$, also observe that[4]

$$\begin{aligned}\mathbf{x}(t) - \mathbf{x}(0) = \int_0^t \mathbf{x}'(s)\, ds &= \int_0^t \Big(A\mathbf{x}(s) + Bu^*(s)\Big)\, ds \\ &= \int_0^t A(\mathbf{x}(s) - \mathbf{x}(0))\, ds + \int_0^t \Big(Ax_0 + Bu^*(s)\Big)\, ds\end{aligned}$$

So by multiply two side with ζ and using (5.4), we have

$$\left\langle \zeta, \int_0^t A(\mathbf{x}(s) - \mathbf{x}(0))\, ds \right\rangle + \left\langle \zeta, \int_0^t \Big(Ax_0 + Bu^*(s)\Big)\, ds \right\rangle \leq 0 \tag{5.5}$$

Our desire is prove $H(x_0, \zeta) \leq 0$, assume the converse hold, then

$$H(x_0, \zeta) = \min_{a\in U}\langle \zeta, Ax_0 + Ba\rangle = \varepsilon > 0$$

since $U = [-1, 1]^m$ is compact. And thus we can assume the minimum hold at $a = \overline{u} \in U$. In this case

$$t\varepsilon \leq \left\langle \zeta, \int_0^t \Big(Ax_0 + Bu^*(s)\Big)\, ds \right\rangle \tag{5.6}$$

[3]That is at every point $x \in \partial\mathcal{R}(T)$, there exists a differentiable map $\alpha : (-\varepsilon, \varepsilon) \longrightarrow \partial\mathcal{R}(T)$ such that $\alpha(0) = x$

[4]In this case, the control u^* is optimal, so by maximum principle it must have finite switching time, denoted by $0 \leq a_1 \leq a_2 \leq \ldots \leq a_n = T$, then on each interval $[a_i, a_{i+1}]$, u^* is a constant, so $\mathbf{x}'(.)$ is continuous on each interval, and the fundamental theorem of calculus can be use here. And we can obtain the this formula straight–forward here.

Now using (5.5) and (5.6) and the fact that $\|\mathbf{x}(s)-\mathbf{x}(0)\| \le Ms$ for all $s \in [0,T]$, we have

$$\begin{aligned} t\varepsilon \le \left\langle \zeta, \int_0^t \Big(Ax_0 + Bu^*(s)\Big)\, ds \right\rangle &\le \left\langle \zeta, \int_0^t A\Big(\mathbf{x}(0)-\mathbf{x}(s)\Big)\, ds \right\rangle \\ &\le \|\zeta\|.\|A\| \int_0^t \|\mathbf{x}(s)-\mathbf{x}(0)\|\, dt \\ &\le \|\zeta\|.\|A\| \int_0^t Ms\, dt = \|\zeta\|.\|A\|.M.\frac{t^2}{2} \end{aligned}$$

From this we have

$$\varepsilon \le \|\zeta\|\|A\|M\frac{t}{2} \qquad \forall\ t \in\ (0,T)$$

It give a contradiction if we let $t \longrightarrow 0$. Thus we conclude that $H(x_0,\zeta) \le 0$. □

5.2 Non-Lipschitz points of the minimum time functions

We begin with the characteristic of non-Lipschitz point.

Proposition 5.7: Let $\Omega \subset \mathbb{R}^n$ be open and $f : \Omega \longrightarrow \mathbb{R}$ is continuous such that epi(f) has positive reach, then f is non-Lipschitz at x, i.e f is strictly continuous at x if and only if there exists a non-zero vector $\zeta \in \mathbb{R}^n$ such that

$$(\zeta,0) \in N^P_{\text{epi}(f)}(x,f(x))$$

Proof. It's well study in theorem 9.13 of [1] that in case epi(f) has positive reach, then f is non-Lipschitz at x if and only if $\partial_\infty f(x)$ contains a non-zero vector $\zeta \in \mathbb{R}^n$, combine this fact with definition 2.25, we can see that

$$\zeta \in \partial_\infty f(x) \Longleftrightarrow (\zeta,0) \in N^P_{\text{epi}(f)}(x,f(x))$$

□

In particular, by theorem 5.2 we obtain

Corollary 5.8: Under the assumption $(\star)$, the minimum time function T is non-Lipschitz at x if and only if there exists a non-zero vector $\zeta \in \mathbb{R}^n$ such that

$$(\zeta,0) \in N^P_{\text{epi}(T)}(x,T(x))$$

Now we prove the main result about non-Lipschitz points of T.

Theorem 5.9: Let $x_0 \neq 0$ in the control system (**??**), and a non-zero vector $\zeta \in \mathbb{R}^n$, then

$$\zeta \in \partial_\infty T(x_0) \quad \text{if and only if} \quad H(x_0,\zeta) = 0 \text{ and } \zeta \in N_{\mathcal{R}(T)}(x_0)$$

Proof. Denote $T = T(x_0)$ as usual.

(i) Assume $\zeta \in \partial_\infty T(x_0)$, i.e $(\zeta, 0) \in N^P_{\text{epi(T)}}(x_0, T)$, equivalently, for any $(z, s) \in \mathbb{R}^n \times \mathbb{R}$ such that $T(z) \leq s$, exists $\sigma = \sigma(\zeta, x_0) > 0$ such that

$$\langle (\zeta, 0), (z, s) - (x_0, T) \rangle \leq \sigma \Big(\|z - x_0\|^2 + |s - T|^2 \Big) \tag{5.7}$$

Take $z \in \mathcal{R}(T)$ and $s = T$, then substituting into above function

$$\langle \zeta, z - x_0 \rangle \leq \sigma \|z - x_0\|^2 \qquad \forall\, z \in \mathcal{R}(T)$$

This equation implies that $\zeta \in N^P_{\mathcal{R}(T)}(x_0)$. But recall that $\mathcal{R}(T)$ is convex, so by using proposition 2.24 we have

$$\zeta \in N^P_{\mathcal{R}(t)}(x_0) = N_{\mathcal{R}(t)}(x_0)$$

Now we will check $H(x_0, \zeta) = 0$. By theorem 5.6, $H(x_0, \zeta) \leq 0$. Since $U = [-1, 1]^m$ is compact, there exists $w \in U$ such that

$$H(x_0, \zeta) = \min_{a \in U} \langle \zeta, Ax_0 + Ba \rangle = \langle \zeta, Ax_0 + Bw \rangle \tag{5.8}$$

Now consider the control system

$$\begin{cases} z'(t) & = -Az(t) - Bw \\ z(0) & = x_0 \end{cases}$$

We observe that the $z'(0) = -(Ax_0 + Bw)$. We can choose time ε so small to ensure that $z(t)$ is not reach to the origin in $[0, 2\varepsilon]$. Fix any $t \in (0, \varepsilon)$, we define

$$\mathfrak{y}(s) = z(t - s) \qquad \forall;\ s \in [0, t]$$

then clearly $\mathfrak{y}(.)$ is the trajectory isn't reach to 0 in $[0, t]$. By Dynamic Programming Principle 2.68, we have

$$T(z(t)) = T(\mathfrak{y}(0)) \leq t + T(\mathfrak{y}(t)) = t + T(z(0)) = t + T(x_0) = t + T$$

Substitute this fact into the condition $\zeta \in \partial^\infty T(x_0)$, i.e (5.7), with $y = z(t)$ and $s = T + t \geq T(z(t))$

$$\langle \zeta, z(t) - z(0) \rangle \leq \sigma \Big(\|z(t) - z(0)\|^2 + t^2 \Big) \leq \sigma(M+1)t^2$$

Thus

$$\left\langle \zeta, \frac{z(t) - z(0)}{t} \right\rangle \leq \sigma(M+1)t$$

Letting $t \longrightarrow 0^+$, we have

$$\langle \zeta, z'(0) \rangle \leq 0 \quad \Longleftrightarrow \quad \langle \zeta, Ax_0 + Bw \rangle \geq 0$$

So we conclude that $H(x_0, \zeta) \geq 0$. Therefore we have $H(x_0, \zeta) = 0$.

(ii) Now assume $H(x_0,\zeta)=0$ and $\zeta\in N_{\mathcal{R}(T)}(x_0)$. We need to show that $(\zeta,0)\in N^P_{\text{epi}}(x_0,T)$, i.e, there exists $\sigma>0$ such that

$$\langle\zeta,z-x_0\rangle\le\sigma\Big(\|z-x\|^2+|\beta-T|^2\Big)\tag{5.9}$$

for all $(z,\beta)\in\text{epi}(T)\cap B((x,T),T)$, i.e $T(z)\le\beta$ and

$$\|z-x\|^2+|\beta-T|^2<T^2\quad\Longrightarrow\quad\beta<2T\tag{5.10}$$

Here we have used the locally definition of proximal cone. Observe that we only need to prove (5.9) when $T(z)=\beta$. There are two cases

- If $T(z)\le T$, then $z\in\mathcal{R}(T)$, since $\zeta\in N_{\mathcal{R}(T)}(x_0)$, we have

$$\langle\zeta,z-x_0\rangle\le 0$$

  because $\mathcal{R}(T)$ is convex, so (5.9) is true.

- If $T<T(z)=\beta$. Let's call the optimal control which steers x_0 to the origin in time T is $u(.):[0,T]\longrightarrow U$. Now by theorem 4.13, x_0 is an optimal point so there exists $x'\in\mathcal{R}$ with $T(x')=\beta$, with the corresponding optimal control $v(.):[0,\beta]\longrightarrow U$ such that

$$\mathbf{x}(\beta-T)=y^{x',v}(\beta-T)=x_0\tag{5.11}$$

  By theorem 4.11, there exists a absolutely continuous function $\psi:[0,\beta]\longrightarrow\mathbb{R}^n$ such that

$$\begin{cases}\psi'(t) &=\psi(t)\cdot A^T\\ \psi(T) &=\zeta\\ \psi(\beta) &\in N_{\mathcal{R}(\beta)}(x')\end{cases}\tag{5.12}$$

  we also know $\psi(t)=e^{-(T-t)A^T}\zeta$, and furthermore it has the property for all $t\in[0,\beta]$

$$H(\mathbf{x}(\beta-t),\psi(t))=\langle\psi(t),A\mathbf{x}(\beta-t)+Bv(\beta-t)\rangle=H(x_0,\zeta)=0\tag{5.13}$$

  On the other hand, for $t\in[0,\beta]$, since (5.10) we have $T<\beta<2T$, so

$$|T-t|=\begin{cases}T-t\le T & \forall\ 0\le t\le T\\ t-T\le\beta-T<T & \forall\ T\le t\le\beta\end{cases}\Longrightarrow|T-t|\le T\qquad\forall\ t\in[0,\beta]$$

  Thus we have

$$\|\psi(t)\|=\|e^{-(T-t)A^T}\zeta\|\le\|\zeta\|\cdot e^{|T-t|\cdot\|A\|}\le\|\zeta\|\cdot e^{T\cdot\|A\|}\le\|\zeta\|\cdot e^{2T\cdot\|A\|}\tag{5.14}$$

  Now we have

$$\langle\zeta,z-x_0\rangle=\langle\psi(\beta),z-x'\rangle+\langle\zeta-\psi(\beta),z-x'\rangle+\langle\zeta,x'-x_0\rangle\tag{5.15}$$

- For the first term, since $\psi(\beta) \in N_{\mathcal{R}(\beta)}(x')$, and $z \in \mathcal{R}(\beta)$, so

$$\langle \psi(\beta), z - x' \rangle \leq 0 \tag{5.16}$$

since $\mathcal{R}(\beta)$ is convex.

- For the second term, by the locally Lipschitz property of the trajectory $y^{x',\nu}(.)$ in theorem 2.71, we have

$$\left\| y^{x',\nu}(\beta - T) - y^{x',\nu}(0) \right\| \leq 2\|B\| e^{2\beta\|A\|}(\beta - T) \leq \|B\| e^{4T\|A\|}(\beta - T)$$

Using this fact, we have

$$\begin{aligned} \|z - x'\| \leq \|z - x_0\| + \|x_0 - x'\| &= \|z - x_0\| + \left\| y^{x',\nu}(\beta - T) - y^{x',\nu}(0) \right\| \\ &\leq \|z - x_0\| + 2\|B\| e^{4T\|A\|} |\beta - T| \end{aligned} \tag{5.17}$$

From the formula of ψ in (5.12) and (5.14) we have

$$\begin{aligned} \|\psi(\beta) - \psi(T)\| = \left\| \int_T^\beta \psi'(s)\, ds \right\| &\leq \|A\| \int_T^\beta \|\psi(s)\|\, ds \leq \|A\| \int_T^\beta \|\zeta\| e^{2T\|A\|}\, ds \\ &= \|A\| \|\zeta\| e^{2T\|A\|} (\beta - T) \end{aligned} \tag{5.18}$$

Apply (5.17) and (5.18) into the second term, we have the estimate

$$\begin{aligned} \langle \zeta - \psi(\beta), z - x' \rangle &\leq \|\Psi(T) - \psi(\beta)\| \cdot \|z - x'\| \\ &\leq \|A\| \|\zeta\| e^{2T\|A\|} |\beta - T| \cdot \left(\|z - x_0\| + 2\|B\| e^{4T\|A\|} |\beta - T| \right) \\ &\leq \|A\| \|\zeta\| e^{2T\|A\|} \|z - x_0\| |\beta - T| + 2\|A\|.\|B\| \|\zeta\| e^{6T\|A\|} |\beta - T|^2 \\ &\leq C_1 \left(\|z - x_0\|^2 + |\beta - T|^2 \right) \end{aligned} \tag{5.19}$$

where C_1 is a constant[5] just depend on T and ζ (and $\|A\|$, $\|B\|$).

- For the third term, we have some observations

  1. Firstly, for any $s \in [0, \beta - T]$, by proposition 2.70 we have

  $$\|\mathbf{x}(s)\| \leq \frac{\|B\|}{\|A\|} e^{(\beta - s)\|A\|} \leq \frac{\|B\|}{\|A\|} e^{\beta\|A\|} \leq \frac{\|B\|}{\|A\|} e^{2T\|A\|}$$

  So using this fact, we have for any $s \in [0, \beta - T]$

  $$\|\mathbf{x}'(s)\| = \|A\mathbf{x}(s) + Bv(s)\| \leq \|A\| \cdot \|\mathbf{x}(s)\| + \|B\| = \|B\|(e^{2T\|A\|} + 1)$$

  2. For any $s \in [0, \beta - T]$, by using (5.14) we have

  $$\begin{aligned} \|\psi(\beta - s) - \psi(T)\| \leq \int_T^{\beta - s} \|\psi'(t)\|\, dt &\leq \|A\| \int_T^{\beta - s} \|\psi(t)\|\, dt \\ &\leq \|A\| \|\zeta\| e^{2T\|A\|} (\beta - T - s) \end{aligned}$$

[5] We have used the AM-GM inequality $xy \leq \frac{x^2 + y^2}{2}$ for $\|z - x_0\|$ and $|\beta - T|$

3. For any $s \in [0, \beta]$, by using (5.13) we have

$$\langle \psi(\beta - s), \mathbf{x}'(s) \rangle = \Big\langle \psi(\beta - s), A\mathbf{x}(s) + Bv(s) \Big\rangle = 0 \tag{5.20}$$

We can summarize these facts by saying that there exists constants C_2, C_3 just depend on $T, \zeta, ||A||, ||B||$ such that for any $s \in [0, \beta - T]$ then

$$||\mathbf{x}'(s)|| \le C_2 \qquad \text{and} \qquad ||\psi(\beta - s) - \psi(T)|| \le C_3(\beta - T - s)$$

So combine these fact we have

$$\begin{aligned}\int_0^{\beta - T} \Big\langle \psi(\beta - s) - \psi(T), \mathbf{x}'(s) \Big\rangle \, ds &\le \int_0^{\beta - T} ||\psi(\beta - s) - \psi(T)|| \cdot ||\mathbf{x}'(s)|| \, ds \\ &\le C_2 C_3 \int_0^{\beta - T} (\beta - T - s) \, ds \\ &= \frac{C_2 C_3}{2} |\beta - T|^2 \end{aligned} \tag{5.21}$$

Now we estimate the third term, observe that

$$\begin{aligned}\langle \zeta, x' - x_0 \rangle = \Big\langle \zeta, \mathbf{x}(0) - \mathbf{x}(\beta - T) \Big\rangle &= -\left\langle \zeta, \int_0^{\beta - T} \mathbf{x}'(s) \, ds \right\rangle \\ &= -\left\langle \psi(\beta - s), \int_0^{\beta - T} \mathbf{x}'(s) \, ds \right\rangle + \left\langle \psi(\beta - s) - \psi(T), \int_0^{\beta - T} \mathbf{x}'(s) \, ds \right\rangle \\ &= -\int_0^{\beta - T} \langle \psi(\beta - s), \mathbf{x}'(s) \rangle \, ds + \int_0^{\beta - T} \Big\langle \psi(\beta - s) - \psi(T), \mathbf{x}'(s) \Big\rangle \, ds \end{aligned}$$

Now from (5.20) and (5.21) we conclude that

$$\langle \zeta, x' - x_0 \rangle \le \frac{C_2 C_3}{2} |\beta - T|^2 \tag{5.22}$$

Putting together (5.16), (5.19) and (5.22) and substitute into (5.15) we conclude that there exists a constant $C > 0$ just depend on $T, \zeta, ||A||, ||B||$ such that

$$\langle \zeta, z - x_0 \rangle \le C \Big(||z - x_0||^2 + |\beta - T|^2 \Big)$$

i.e, (5.9) is true.

The proof is complete. $\square$

By this result, we will establish the detailed structure of the set of all non-Lipschitz points S_T^∞ of T. A point $x_0 \in S_T^\infty$ if and only if there exists $0 \neq \zeta \in N_{\mathcal{R}(T(x_0))}(x_0)$ such that

$$H(x_0, \zeta) = 0$$

First we will establish the explicit formula for the minimized Hamiltonian $H(x, p)$

Proposition 5.10 (The explicit formula of minimized Hamiltonian): Assume $x_0 \in \partial\mathcal{R}(T)$, let $u^* \in \mathcal{U}_{ad}$ be the optimal control which is respect to x_0, then x can be presented as

$$x_0 = -\int_0^T e^{-sA}Bu^*(s)\,ds \tag{5.23}$$

Then

$$H(x_0,\zeta) = -\sum_{i=1}^m \left|\left\langle \zeta, e^{-TA}b_i\right\rangle\right|$$

Proof. By maximum principle 4.8, the control u^* is unique (in sense almost every where) and there exists $0 \neq \zeta \in N_{\mathcal{R}(T)}$ such that

$$u^*(t) = \begin{pmatrix} u_1^*(t) \\ \vdots \\ u_m^*(t) \end{pmatrix} \qquad \text{and} \qquad u_i^*(t) = -\text{sign}\left\langle \zeta, e^{-tA}b_i\right\rangle$$

for $t \in [0,T]$. Combine with (5.23) we have

$$x_0 = \sum_{i=1}^m \int_0^T e^{-sA}b_i \text{sign}\,\langle \zeta, e^{-sA}b_i\rangle\,ds$$

and thus

$$\langle \zeta, Ax_0\rangle = \sum_{i=1}^m \int_0^T \left\langle \zeta, Ae^{-sA}b_i \text{sign}\,\langle \zeta, e^{-sA}b_i\rangle\right\rangle ds$$

So by definition of the minimized Hamiltonian, we have

$$\begin{aligned} H(x_0,\zeta) &= \langle \zeta, Ax_0\rangle + \min_{u\in U}\langle \zeta, Bu\rangle = \langle \zeta, Ax_0\rangle + \min_{u_i\in[-1,1]}\sum_{i=1}^m \langle \zeta, b_iu_i\rangle = \langle \zeta, Ax_0\rangle - \sum_{i=1}^m |\langle \zeta, b_i\rangle| \\ &= \sum_{i=1}^m\left(\int_0^T \left\langle \zeta, Ae^{-sA}b_i \text{ sign}\,\langle \zeta, e^{-sA}b_i\rangle\right\rangle ds - |\langle \zeta, b_i\rangle|\right) = \sum_{i=1}^m H_i(x_0,\zeta) \end{aligned} \tag{5.24}$$

where

$$H_i(x_0,\zeta) = \int_0^T \left\langle \zeta, Ae^{-sA}b_i \text{ sign}\,\langle \zeta, e^{-sA}b_i\rangle\right\rangle ds - |\langle \zeta, b_i\rangle|$$

and

$$\begin{aligned} g_i : [0,T] &\longrightarrow \mathbb{R} \\ s &\longmapsto \langle \zeta, e^{-sA}b_i\rangle \end{aligned}$$

then similarly to the proof of proposition 4.2, g_i can only equal to zero in finite point in $[0,T]$, we can assume $0 \le s_1 \le s_2 \le \ldots \le s_{k-1} \le s_k \le T$ and g_i only vanish at $s_1, s_2, \ldots, s_{k-1}, s_k$. Then on each $[s_i, s_{i+1}]$, sign $g_i(s)$ is a constant 1 or -1. Assume

$$\text{sign}(g_i(s)) = \alpha \quad \text{on } [s_1,s_2] \quad \Longrightarrow \quad \text{sign}(g_i(s)) = \alpha(-1)^{j-1} \quad \text{on } [s_j, s_{j+1}]$$

Furthermore

$$\begin{aligned} g_i' : [0,T] &\longrightarrow \mathbb{R} \\ s &\longmapsto g_i'(s) = -\langle \zeta, Ae^{-sA} b_i \rangle \end{aligned}$$

By this notion, $H_i(x_0,\zeta)$ can be rewrite as

$$\begin{aligned} H_i(x_0,\zeta) &= -\int_0^T g_i'(s) \cdot \mathrm{sign}\,(g_i(s))\,ds - |\langle \zeta, b_i \rangle| \\ &= -\int_0^{s_1} g_i'(s)\mathrm{sign}(g_i(s))\,ds - \sum_{j=1}^{k-1} (-1)^{j-1}\alpha \int_{s_j}^{s_{j+1}} g_i'(s)\,ds - \int_{s_k}^T g_i'(s)\mathrm{sign}(g_i(s))\,ds - |\langle \zeta, b_i \rangle| \\ &= -\int_0^{s_1} g_i'(s)\mathrm{sign}(g_i(s))\,ds + \alpha \sum_{j=0}^{k-1} (-1)^j \Big[g_i(s_{j+1}) - g(s_j) \Big] - \int_{s_k}^T g'(s)\mathrm{sign}(g_i(s))\,ds - |\langle \zeta, b_i \rangle| \\ &= -\int_0^{s_1} g_i'(s)\mathrm{sign}(g_i(s))\,ds - \int_{s_k}^T g_i'(s)\mathrm{sign}(g(s))\,ds - |\langle \zeta, b_i \rangle| \end{aligned}$$

since $g_i(s_j) = 0$ for all $j = 1, 2 \ldots, k$.

- If $g_i(0) \neq 0$ and $g_i(T) \neq 0$, then $\mathrm{sign} g(s) = \mathrm{sign} g(0)$ for all $s \in [0, s_1)$, similarly $\mathrm{sign} g(s) = \mathrm{sign} g(T)$ for all $s \in (s_k, T]$, therefore

$$H_i(x_0,\zeta) = g_i(0)\mathrm{sign}(g_i(0)) - g_i(T)\mathrm{sign}(g_i(T)) - g_i(0)\mathrm{sign}(g_i(0)) = -|g_i(T)|$$

- If $g_i(0) = 0$ and $g_i(T) \neq 0$, then $\mathrm{sign} g(s) = \mathrm{sign} g(T)$ for all $s \in (s_k, T]$, therefore

$$H_i(x_0,\zeta) = -g_i(T)\mathrm{sign}(g_i(T)) = -|g_i(T)|$$

- If $g_i(0) \neq 0$ and $g_i(T) = 0$, then $\mathrm{sign} g(s) = \mathrm{sign} g(0)$ for all $s \in [0, s_1)$, therefore

$$H_i(x_0,\zeta) = g_i(0)\mathrm{sign}(g_i(0)) - g_i(0)\mathrm{sign}(g_i(0)) = 0 = -|g_i(T)|$$

- If $g_i(0) = 0$ and $g_i(T) = 0$, then $\langle \zeta, b_i \rangle = g_i(0) = 0 = \langle \zeta, e^{-TA} b_i \rangle$, so

$$H_i(x_0,\zeta) = 0 = -|g_i(T)|$$

Therefore in general, we always have

$$H_i(x_0,\zeta) = -|g_i(T)| = -\left| \left\langle \zeta, e^{-TA} b_i \right\rangle \right|$$

and now combine this fact with (5.24) we conclude that

$$H(x_0,\zeta) = \sum_{i=1}^m H_i(x_0,\zeta) = -\sum_{i=1}^m \left| \left\langle \zeta, e^{-TA} b_i \right\rangle \right|$$

and the proof is complete. □

With this formula of H, we conclude that

Theorem 5.11: Assume $B = [b_1, b_2, \dots, b_m]$ are m columns, under the assumption $(\star)$, setting

$$\mathcal{S} = \left\{ x \in \mathbb{R}^n : \begin{array}{c} \exists\, T > 0 \\ \exists\, \zeta \in \mathbb{R}^n \backslash \{0\} \end{array} \text{ s.t } \langle \zeta, e^{-TA} b_i \rangle = 0 \;\forall\, i = \overline{1,m} \text{ and } x = \sum_{i=1}^{m} \int_0^T e^{-sA} b_i \operatorname{sign}\langle \zeta, e^{-sA} b_i \rangle \, ds \right\}$$

$$\mathcal{S}' = \left\{ x \in \mathbb{R}^n, \exists\, T > 0 \text{ s.t } x \in \partial \mathcal{R}(T) \text{ and } 0 \neq \zeta \in N_{\mathcal{R}(T)}(x) \text{ for which } H(x,\zeta) = 0 \right\}$$

then $S_T^{\infty} = \mathcal{S} = \mathcal{S}'$. I.e, T is non-Lipschitz at x if and only if $x \in \mathcal{S}$.

Proof. Let $x \neq 0$ be a non-Lipschitz point of T, set $T = T(x) > 0$, then clearly $x \in \partial \mathcal{R}(T)$. By corollary 5.8 there exists a non–zero vector $\zeta \in \mathbb{R}^n$ such that

$$(\zeta, 0) \in N^P_{\mathrm{epi}(T)}(x, T) \Longleftrightarrow \langle \zeta, y - x \rangle \leq \sigma(\zeta, x)\Big(\|y - x\|^2 + |\beta - T|^2 \Big)$$

for all $(y, \beta) \in \mathrm{epi}(T)$, where $\sigma(\zeta, x) > 0$. Let $\beta = T$, we obtain

$$(\zeta, 0) \in N^P_{\mathrm{epi}(T)}(x, T) \Longleftrightarrow \langle \zeta, y - x \rangle \leq \sigma(\zeta, x) \|y - x\|^2$$

i.e, $\zeta \in N_{\mathcal{R}(T)}(x)$. Now by theorem 5.9 we obtain $H(x, \zeta) = 0$. So

$$S_T^{\infty} \subseteq \mathcal{S}' \tag{5.25}$$

Now let u^* is the optimal control which steers x to the origin in time T. By corollary 4.8, u is unique and furthermore, it has the form

$$u(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_m(t) \end{pmatrix} \qquad \text{where} \qquad u_i(t) = -\operatorname{sign}\langle \zeta, e^{-tA} b_i \rangle$$

Consequently,

$$x = -\int_0^T e^{-sA} B u(s)\, ds = \sum_{i=1}^{m} \int_0^T e^{-sA} b_i \operatorname{sign}\langle \zeta, e^{-sA} b_i \rangle \, ds$$

On the other hand, by proposition 5.10, we have

$$H(x, \zeta) = -\sum_{i=1}^{m} \left| \left\langle \zeta . e^{-TA} b_i \right\rangle \right| = 0 \Longrightarrow \langle \zeta, e^{-TA} b_i \rangle = 0 \;\forall\, i = \overline{1,m}$$

Therefore we have

$$\mathcal{S}' \subseteq \mathcal{S} \tag{5.26}$$

Now if $x \in \mathcal{S}$, then there exists $T > 0$ and $0 \neq \zeta \in \mathbb{R}^n$ such that $\langle \zeta, e^{-TA} b_i \rangle = 0$ for all $i = \overline{1,m}$, and

$$x = \sum_{i=1}^{m} \int_0^T e^{-sA} b_i \operatorname{sign}\langle \zeta, e^{-sA} b_i \rangle \, ds$$

then the control u which steer x to the origin in time T has form

$$u(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_m(t) \end{pmatrix} \qquad \text{where} \qquad u_i(t) = -\text{sign}\langle \zeta, e^{-tA} b_i \rangle$$

By theorem 4.9, the control u is optimal, $T(x) = T$, and $\zeta \in N_{\mathcal{R}}(T)(x)$. By proposition 5.10, $\langle \zeta, e^{-TA} b_i \rangle = 0$ for all $i = \overline{1, m}$ implies $H(x, \zeta) = 0$. So by theorem 5.9 we conclude that

$$\zeta \in \partial^\infty T(x)$$

so by corollary 5.8 we have T is non–Lipschitz at x, i.e

$$\mathcal{S} \subseteq S_T^\infty \tag{5.27}$$

Finally, combine (5.25), (5.26) and (5.27) we obtain $\mathcal{S} = \mathcal{S}' = S_T^\infty$. □

Now recall theorem 4.11, we obtain the following result

Theorem 5.12: Under the assumption ($\star$). If $x_0 \in \partial\mathcal{R}(T)$, let $u^* \in \mathcal{U}_{ad}$ is the corresponding optimal control and $\mathbf{x}(.) = y^{x_0, u^*}(.)$ is the corresponding trajectory. If $x_0 \in \mathcal{S}$, i.e T is non-Lipschitz at x_0, then $\mathbf{x}(T - t) \in \mathcal{S}$ for all $t \in [0, T]$.

Proof. Since T is non-Lipschitz at x_0, there exists a non-zero vector $\zeta \in N_{\mathcal{R}(T)}(x_0)$ such that $H(x_0, \zeta) = 0$. Let $\psi : [0, T] \longrightarrow \mathbb{R}^n$ is the function in theorem 4.11 such that $\psi(T) = \zeta$, then

$$H(\mathbf{x}(T - t), \psi(t)) = H(x_0, \zeta) = 0 \qquad \forall\ t \in [0, T]$$

Furthermore, theorem 4.11 states that

$$0 \neq \psi(t) \in N_{\mathcal{R}(t)}(\mathbf{x}(T - t)) \qquad \forall\ t \in (0, T)$$

So by theorem 5.9 and 5.11 we obtain $\mathbf{x}(T - t) \in \mathcal{S}$ for all $t \in [0, T]$. □

Similarly, under the assumption ($\star$), by using theorem 4.13 we obtain the following result.

Theorem 5.13: Under the assumption ($\star$), if $x_\alpha \in \partial\mathcal{R}(\alpha)$, assume T is non–Lipschitz at x_α, then by theorem 5.9, there exists a non-zero vector $\zeta \in N_{\mathcal{R}(\alpha)}(x_\alpha)$ such that $H(x_\alpha, \zeta) = 0$. By theorem 4.13, setting

$$\begin{aligned} \Psi : [0, \infty) &\longrightarrow \mathbb{R}^n \\ t &\longmapsto \langle \zeta, e^{(\alpha - t)A} b_i \rangle \end{aligned}$$

and for any $\beta > \alpha$, setting the control $u_\beta : [0, \beta] \longrightarrow \mathbb{R}^m$ by

$$u_\beta(t) = \begin{pmatrix} v_1(t) \\ \vdots \\ v_m(t) \end{pmatrix} \quad \text{where} \quad v_i(t) = -\text{sign}\langle \psi(\beta), e^{-tA} b_i \rangle$$

and

$$x_\beta = -\int_0^\beta e^{-tA} B u_\beta(t)\, dt$$

Then $H(x_\beta, \Psi(\beta)) = 0$ for all $\beta > \alpha$. In particular, $x_\beta \in \mathcal{S}$ for all $\beta > \alpha$.

Proof. By theorem 4.13, $x_\beta \in \partial\mathcal{R}(\beta)$, and furthermore

$$y^{x_\beta,u_\beta}(\beta-\alpha)=x_\alpha$$

Now doing similarly to theorem 5.12, we obtain $\Psi(\beta)\in N_{\mathcal{R}(\beta)}(x_\beta)$ and by theorem 4.11, we have

$$H(x_\beta,\Psi(\beta))=H(y^{x_\beta,u_\beta}(\beta-t),\Psi(t))=H(x_\alpha,\zeta)=0$$

Thus by theorem 5.9 and 5.11, we conclude that $x_\beta\in\mathcal{S}$. □

Above theorems give us a method to calculate the set $\mathcal{S}$ in some cases. That is we only need to calculate some point of $x_0\in\mathcal{S}$ that lie in particular $\partial\mathcal{R}(T)$, then using the extended trajectory inside and outside $\mathcal{R}(T)$, we easily obtain all non-Lipschitz points of T that lie in a trajectory pass through x_0. But some simple case, we can doing straight-forward, as in the example Rocket car problem following.

5.2.1 Application to the Rocket Rail road car problem

Recall the Rocker Rail Road car problem,

$$\begin{cases} z'(t) &= A\cdot z(t)+b\cdot u(t)\\ z(0) &= (x_0,v_0)^T\end{cases}$$

where $u(t)\in U=[-1,1]$ and

$$z(t)=\begin{pmatrix} x(t)\\ v(t)\end{pmatrix}\qquad A=\begin{pmatrix}0&1\\0&0\end{pmatrix}\qquad b=\begin{pmatrix}0\\1\end{pmatrix}$$

We have know that in this problem the assumption ($\star$) hold, so $\mathcal{R}=\mathbb{R}^n$. Also we have

$$e^{-tA}=\begin{pmatrix}1&-t\\0&1\end{pmatrix}\qquad\forall\ t>0$$

Now using theorem 5.11, on the boundary $\mathcal{R}(t)$, we find x_0^t such that there exists $\zeta\in N_{\mathcal{R}(t)}(x_0^t)$ and $\langle\zeta,e^{-tA}b\rangle=0$. Assume $\zeta=(\zeta_1,\zeta_2)$ then

$$\langle\zeta,e^{-tA}b\rangle=\begin{pmatrix}\zeta_1&\zeta_2\end{pmatrix}\cdot\begin{pmatrix}1&-t\\0&1\end{pmatrix}\cdot\begin{pmatrix}0\\1\end{pmatrix}=-t\zeta_1+\zeta_2=0$$

We can choose the $\zeta=(1,t)$ and $\zeta=(-1,-t)$ in this case, note that the difference on $\|\zeta\|$ doesn't have any effect.

- Case $\zeta=(1,t)$. Now we calculate the optimal control by

  $$u(s)=-\text{sign}\ \langle\zeta,e^{-sA}b\rangle=-\text{sign}(-s+t)=-1$$

  since $s\in[0,t]$. So, the corresponding point x_0^t is

  $$x_0^t=\int_0^t e^{-sA}b\,\text{sign}\langle\zeta,e^{-sA}b\rangle\ ds=\int_0^t\begin{pmatrix}s\\-1\end{pmatrix}ds=\begin{pmatrix}\frac{t^2}{2}\\-t\end{pmatrix}$$

  This is just the curve $(S_1): x=\frac{v^2}{2}, v<0$.

- Case $\zeta = (-1, -t)$. Now we calculate the optimal control by

$$u(s) = -\text{sign}\,\langle \zeta, e^{-sA} b \rangle = -\text{sign}(s-t) = 1$$

since $s \in [0, t]$. So, the corresponding point x_0^t is

$$x_0^t = \int_0^t e^{-sA} b \,\text{sign} \langle \zeta, e^{-sA} b \rangle \, ds = \int_0^t \begin{pmatrix} -s \\ 1 \end{pmatrix} ds = \begin{pmatrix} -\frac{t^2}{2} \\ t \end{pmatrix}$$

This is just the curve $(S_2): x = -\frac{v^2}{2}, v > 0$.

So we conclude that in this problem, the set of all non–Lipschitz points of T is

$$\mathcal{S} = S_1 \cup S_2 \cup \{0\}$$

since $\mathcal{S}$ is closed.

5.2.2 Application to the Harmonic Oscillator problem

Recall the Harmonic Oscillator problem

$$z'(t) = \begin{pmatrix} v(t) \\ v'(t) \end{pmatrix} \qquad z(0) = \begin{pmatrix} x_0 \\ v_0 \end{pmatrix} \Longrightarrow \begin{cases} z'(t) & = A \cdot z(t) + b \cdot u(t) \\ z(0) & = (x_0, v_0)^T = z_0 \end{cases}$$

where $u(t) \in U = [-1, 1]$ and

$$z(t) = \begin{pmatrix} x(t) \\ v(t) \end{pmatrix} \qquad A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Under the assumption $(\star)$, $\mathcal{R} = \mathbb{R}^n$ and

$$e^{-tA} = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \qquad \forall\ t > 0$$

Now using theorem 5.11, on the boundary $\mathcal{R}(t)$, we find x_0^t such that there exists $\zeta \in N_{\mathcal{R}(t)}(x_0^t)$ and $\langle \zeta, e^{-tA} b \rangle = 0$. Assume $\zeta = (\zeta_1, \zeta_2)$ then

$$\langle \zeta, e^{-tA} b \rangle = \begin{pmatrix} \zeta_1 & \zeta_2 \end{pmatrix} \cdot \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -\sin s \cdot \zeta + \cos s \cdot \zeta_2 = 0 \tag{5.28}$$

We can choose the $\|\zeta\| = 1$, and assume $\zeta = (\zeta_1, \zeta_2) = (-\cos\delta, \sin\delta)$, where $0 \le \delta < 2\pi$. Then (5.28) implies

$$\sin(t + \delta) = 0$$

And thus, we must have $\delta = -t + k\pi$, where $k \in \mathbb{Z}$, so

$$0 \le \delta < 2\pi \Longrightarrow \frac{t}{\pi} \le k < \frac{t}{\pi} + 2 \Longrightarrow k = \left\lceil \frac{t}{\pi} \right\rceil \quad \text{or} \quad \left\lceil \frac{t}{\pi} \right\rceil + 1$$

and thus

$$\zeta^1 = (\zeta_1, \zeta_2) = (-\cos\delta, \sin\delta) = \left(-\cos\left(-t + \left\lceil \frac{t}{\pi} \right\rceil \pi\right), \sin\left(-t + \left\lceil \frac{t}{\pi} \right\rceil \pi\right)\right)$$

or

$$\begin{aligned}\zeta^2 = (\zeta_1, \zeta_2) = (-\cos\delta, \sin\delta) &= \left(-\cos\left(-t + \pi + \left\lceil \frac{t}{\pi} \right\rceil \pi\right), \sin\left(-t + \pi + \left\lceil \frac{t}{\pi} \right\rceil \pi\right)\right) \\ &= \left(\cos\left(-t + \left\lceil \frac{t}{\pi} \right\rceil \pi\right), -\sin\left(-t + \left\lceil \frac{t}{\pi} \right\rceil \pi\right)\right) = -\zeta^1\end{aligned}$$

This fact is compatible with the fact that $\mathcal{R}(t)$ is symmetric through the origin. So we just need to find the trajectory with ζ^1, then the rest is the symmetric part. Now we using this fact to calculate the corresponding optimal control, for $s \in [0, t]$, we have

$$u^1(s) = -\text{sign}\langle \zeta^1, e^{-sA} b \rangle = -\text{sign } \sin(\delta + s) = -\text{sign} \sin\left(s - t + \left\lceil \frac{t}{\pi} \right\rceil \pi\right)$$

Now using this fact to calculate $z(t)$ given by

$$z^1(t) = -\int_0^t e^{-sA} b u^1(s)\, ds = \int_0^t \begin{pmatrix} -\sin s \\ \cos s \end{pmatrix} \text{sign } \sin\left(s - t + \left\lceil \frac{t}{\pi} \right\rceil \pi\right) ds$$

Using this formula, we can draw $\mathcal{S}$ as follow

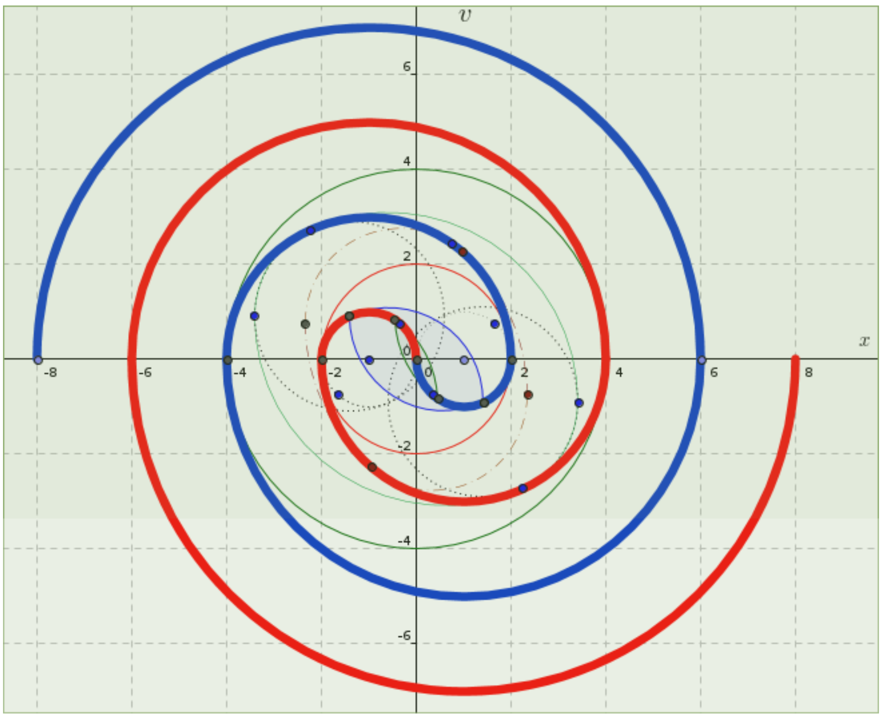


Figure 30: Set of non-Lipschitz points S_T^∞ in Harmonic Oscillator problem

As we can see, $\mathcal{S}$ consists of two curves, which are symmetric through the origin. Furthermore, these curves are tangent to the boundary $\partial\mathcal{R}(t)$ for any $t > 0$.

References

[1] R. T. Rockafellar & R. J-B. Wets, Variational Analysis, Springer, Berlin (1998).

[2] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, P. R. Wolenski, Nonsmooth Analysis and Control Theory, Springer, New York (1998).

[3] G. Colombo, A. Marigonda & P. R. Wolenski, Some new regularity properties for the minimal time function, SIAM J. Control Optim. 44 (2006), 2285–2299.

[4] Khai T. Nguyen, Giovanni Colombo, Luong V. Nguyen, Non-Lipschitz points and the SBV regularity of the minimum time function, Calc. Var. Partial Differential Equations (2013).

[5] Giovanni Colombo, Khai T. Nguyen, On the structure of the minimum time function, SIAM J. Control Optim 48 (2010), no. 7, 4776–4814.

[6] Giovanni Colombo, Khai T. Nguyen, On the minimum time function around the origin, Mathematical Control and Related Fields 3 (2013), no. 1, 51–82.

[7] G. Colombo & A. Marigonda, Differentiability properties for a class of non-convex functions, Calc. Var. Partial Differential Equations 25 (2006), 1–31.

[8] Brezis, Haim , Functional Analysis, Sobolev Spaces and Partial Differential Equations (2011), Springer.

[9] Lawrence C. Evans, Partial Differential Equations: Second Edition, Graduate Studies in Mathematics, Volume 19, AMS.

[10] Gerald B. Folland, Real Analysis: Modern Techniques and Their Applications, Second Edittion, April 7, 1999, Pure and Applied Mathematics, A Wiley-Interscience Series of Texts, Monographs, and Tracts.

[11] P. Cannarsa & C. Sinestrari, Convexity properties of the minimum time function, Calc. Var. Partial Differential Equations 3 (1995), 273-298.

[12] P. Cannarsa, C. Sinestrari, Semiconcave Functions, Hamilton-Jacobi Equations, and Optimal Control, Birkhäuser

[13] H. Federer, Curvature measures, Trans. Amer. Math. Soc 93 (1959), 418–491.

[14] Clarke F.H., Steren R.J., Wolenski P.R., Proximal smoothness and the lower-C^2 property, J. Convex Anal. 2 (1995), 117–144.

[15] A. Marigonda, Khai T. Nguyen, D. Vittone, Some regularity results for a class of upper semicontinuous functions, Indiana Univ. Math. J., in print.

[16] Khai T. Nguyen, Hypographs satisfying an external sphere condition and the regularity of the minimum time function, J. Math. Anal. Appl. 372 (2010), 611–628.

[17] R. Vinter, Optimal control, Birkhäuser, Boston (2000)